

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run On: July 3, 2003, 02:31:01 ; Search time 4275.04 seconds
(without alignments)
11062.367 Million cell updates/sec

Title: US-10-004-219B-2
Perfect score: 1625
Sequence: 1 gcttccagctggtggtga.....ctgcaataaaatcagcagtc 1625

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rod.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match	Length	DB	ID	Description
1	1625	100.0	1625	9	AF290004	AF290004	Homo sapi
2	1266.2	77.9	1369	6	AX405989	AX405989	Sequence
3	1260.2	77.6	1354	9	AB025008	AB025008	Homo sapi
4	1095.2	67.4	1529	4	AB051629	AB051629	Bos tauru
5	1046	64.4	4250	9	AK098814	AK098814	Homo sapi
6	1045.6	64.3	1557	10	BC034548	BC034548	Mus muscu
7	1042.2	64.1	1538	10	BC011134	BC011134	Mus muscu
8	1041.8	64.1	1526	10	AF154571	AF154571	Mus muscu
9	1041.4	64.1	1530	10	AF290003	AF290003	Mus muscu
10	1037.2	63.8	1188	9	AB025009	AB025009	Homo sapi
11	701.8	43.2	1525	10	M94584	M94584	Mus musculu
12	697.2	42.9	1506	10	D87757	D87757	Mus musculu
13	689.4	42.4	1209	10	AY049765	AY049765	Mus muscu
14	687.8	42.3	1209	10	AY065557	AY065557	Mus muscu
15	483	29.7	1656	6	AR206042	AR206042	Sequence
16	483	29.7	1656	6	AR212139	AR212139	Sequence
17	483	29.7	1656	6	AX108752	AX108752	Sequence
18	481.4	29.6	1633	9	HSU29615	U29615	Human chito
19	481.4	29.6	1643	6	AR172107	AR172107	Sequence
20	481	29.6	1636	6	AR206041	AR206041	Sequence
21	481	29.6	1636	6	AR212138	AR212138	Sequence
22	481	29.6	1636	6	AX108750	AX108750	Sequence
23	477.2	29.4	1713	6	AR172108	AR172108	Sequence
24	476.8	29.3	1710	9	HSU62662	U62662	Homo sapien
25	457	28.1	994	10	MMU56900	U56900	Mus musculu
26	427	26.3	1562	5	BJA345054	AJ345054	Bufo japo
27	389.4	24.0	1599	9	AK055165	AK055165	Homo sapi
28	369	22.7	1433	6	AR042834	AR042834	Sequence
29	369	22.7	1434	9	HSU58514	U58514	Human chiti
30	367.4	22.6	1418	9	HSU49835	U49835	Human YKL-3
31	367.4	22.6	1449	9	BC011460	BC011460	Homo sapi
32	366	22.5	1500	9	HSU58515	U58515	Human chiti
33	360	22.2	2237	9	MMU87259	U87259	Macaca mula
34	356.8	22.0	1406	9	AF484550	AF484550	Macaca ra
35	354.8	21.8	1526	6	AR042835	AR042835	Sequence
36	354.2	21.8	2198	9	HSU09550	U09550	Human ovidu
37	354.2	21.8	2228	9	BABED0SG	M5903	Papio hamad
38	352.4	21.7	2525	10	MUSOGP	D32137	Mouse mRNa
39	350.8	21.6	2504	6	E09046	E09046	cDNA encodi
40	350	21.5	2353	10	MAU15048	U15048	Mesocricetu
41	350	21.5	2387	10	HAM036	D32218	Hamster mRN
42	349.4	21.5	1149	6	E01501	E01501	cDNA encodi
43	349.4	21.5	1741	9	HUMH33	M80927	Human glyco
44	348.4	21.4	91569	9	AL356387	AL356387	Human DNA
45	348.2	21.4	2034	4	OAU16719	U16719	Ovis aries

ALIGNMENTS

RESULT 1
AF290004
LOCUS Homo sapiens acidic mammalian chitinase precursor, mRNA, complete cds.
DEFINITION AF290004.1 GI:12597292
ACCESSION AF290004
VERSION AF290004.1
KEYWORDS
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1625)
AUTHORS Boot,R.G., Blommaert,E.F., Swart,E., Ghausharali-van der Vlugt,K.,
Bijl,N., Moe,C., Place,A. and Aerts,J.M.

IDENTIFICATION OF A NOVEL ACIDIC MAMMALIAN CHITINASE DISTINCT FROM CHITOTRIOSIDASE

J. Biol. Chem. 276 (9), 6770-6778 (2001) *March, 2001*

21125893

PUBMED 11085997

2 (bases 1 to 1625)

Boot, R.G., Verhoek, M., Swart, E. and Aerts, J.M.

Direct Submission

TITLE

JOURNAL

Submitted (26-JUL-2000) Dept. of Biochemistry, Academic Medical Center, University of Amsterdam, Meibergdreef 15, Amsterdam 1105 Az, The Netherlands

FEATURES

Location/Qualifiers

source

1..1625

/organism="Homo sapiens"

/db_xref="taxon:9606"

/tissue_type="stomach; lung"

104..1534

/EC_number="3.2.1.14"

/note="chitinase precursor; AMCase; chitinase activity demonstrated by expression in COS cells; similar to sequences presented in GenBank Accession Numbers AB025008 and AB025009"

/codon_start=1

/product="acidic mammalian chitinase precursor"

/protein_id="AAG60019.1"

/db_xref="GI:1259793"

/translation="MTKLILTLGLVLLILQLGSAYOLTCYETNWAQVRPLGIREMPFNIDPCLATHLIYAFAGRONNEITIEWNDVLYQAFNGKKNKNSQPKDKHLTVLVLTAPFTAMVSPENRQFTITSVIFKRLQTFDGLDFWEIYRSGSPQDKHLTVLVLEMYRQAEQAKQINKPMLWVAAGAIVSNISQSYEIPQLSDYLDYIHMTYDLHGSVEYGTENSPYKPYDTSQSNAYLVNMYNWKONGAPAEKLIYGFPTYGHNFIILSNSTNGIGATPSGAGPAGYAKESGIWWEIETFLKNGATQGWDAQPEVYPAQGVQWASGDYNIKFSDIKAOWLKINKFGAMVAIIDLDDTGTFCNQCKFPLISTLKKALGQSSCTAPAPQIEPITAAPSGSNGSGSSSGSGSGGFCFAVRANGLYPVANNRNFVHGVNGVYQONCOAGLVFDTSCDCNWA"

sig_peptide 104..166

BASE COUNT 387 a 468 c 399 g 371 t

ORIGIN

Query Match 100.0%; Score 1625; DB 9; Length 1625;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1625; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 1 GCTTTCAGTCTGTTGGTGAATCCTTCATAGTCTGAAGCCCTTTGGTAACACAGAATC 60

DB 1 GCTTTCAGTCTGTTGGTGAATCCTTCATAGTCTGAAGCCCTTTGGTAACACAGAATC 60

QY 61 AGAACATATAAAAAGCTCTGGCGACATGGTCTGACTGCAACCATGACAAAAGCTTATTCT 120

DB 61 AGAACATATAAAAAGCTCTGGCGACATGGTCTGACTGCAACCATGACAAAAGCTTATTCT 120

QY 121 CCTCACAGGTCTTGTCTCTTATCTAGTGAATTTGCAGCTCGGCTCTGCCTACCAAGTGCATG 180

DB 121 CCTCACAGGTCTTGTCTCTTATCTAGTGAATTTGCAGCTCGGCTCTGCCTACCAAGTGCATG 180

QY 181 CTACTTACCACACTGGCGCCAGTACCGGCCAGGCGCTGGGGCGCTTCATGCTCTGACAAAT 240

DB 181 CTACTTACCACACTGGCGCCAGTACCGGCCAGGCGCTGGGGCGCTTCATGCTCTGACAAAT 240

QY 241 CGACCCCTGCCTCTGTATCCCAACCTTGATCTACGCCCTTTGCTGGGAGGCAAGACAACGAGAT 300

DB 241 CGACCCCTGCCTCTGTATCCCAACCTTGATCTACGCCCTTTGCTGGGAGGCAAGACAACGAGAT 300

QY 301 CACCACCATCGAATGGAACGATGTGACTCTCTACCAAGCTTTCATAGCCCTGAAAAATAA 360

DB 301 CACCACCATCGAATGGAACGATGTGACTCTCTACCAAGCTTTCATAGCCCTGAAAAATAA 360

QY 361 GAACAGCCAGCTGAAAACTCTCCCTGGCCATTGGAGCTGGAACTTCGGAGCTGCCCTTT 420

DB 361 GAACAGCCAGCTGAAAACTCTCCCTGGCCATTGGAGCTGGAACTTCGGAGCTGCCCTTT 420

QY 421 CACTGCCATGGTTTCTACTCCTGAGAACCCGACGACTTTTCATCACTTCAGTCATCAAAAT 480


```
QY 1561 GTTCAGTCTCTTTTGGCTTAGGACATGTTGCCCTTACCTAAAGTCTCTGCAATAAAATCAG 1620
|||||
Db 1561 GTTCAGTCTCTTTTGGCTTAGGACATGTTGCCCTTACCTAAAGTCTCTGCAATAAAATCAG 1620

QY 1621 CAGTC 1625
|||||
Db 1621 CAGTC 1625

RESULT 2
LOCUS Ax405989 1369 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 404 from Patent WO022660.
ACCESSION Ax405989
VERSION Ax405989.1 GI:21439417
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1. Tang, Y.T., Liu, C., Zhou, P., Asundi, V., Zhang, J., Zhao, Q.A., Ren, F.,
Xue, A.J., Yang, Y., Wehrman, T. and Drmanac, R.T.
Novel nucleic acids and polypeptides
Patent: WO 022660-A 404 21-MAR-2002;
HYSEQ, INC. (US)
FEATURES
source Location/Qualifiers
1..1369
/organism="Homo sapiens"
/db_xref="taxon:9606"
154..1260
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD35096.1"
/db_xref="GI:21439418"
/translation="MVSTPENRQTEITSVIKFLRQYEFDFDWEYFSGSRSPQDK
HLFTVLVQEMREAFQEQAKQINKPRLMTAAVAAGISNIQSGYEIPQLSQYLDYIHYM
TYDLHSGMEGTGENSPLYKYPTDTSNAYLVNDVYMWKNDGAPAEKLIIVGFETYG
HNFILSNPTGIGAPTSGAGPAGYKESGIWAYEICTFLKNGATGWDAPQVPI
AYQNVGVYDNIKSFIDIAQWLKKNKFGAMVWALDDDFGTGFCNKGKFLPILTLK
ALGLQASCTAPAPIPIETPAAPSGNGSGSSSGSGSGSGSGSGSGSGSGSGSGSGSG
NNRNFHCVNGVYTYQNCQAGLVFTTSCDCNNWA"
BASE COUNT 327 a 393 c 341 g 308 t
ORIGIN

Query Match 77.9%; Score 1266.2; DB 6; Length 1369;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 1274; Conservative. 0; Mismatches 13; Indels 0; Gaps 0;

QY 339 CTTTCAATGGCTGAAATAAGAACAGCCAGCTGAAACTCTCTGGCCATTGGAGGCT 398
|||||
Db 65 CTTTACTGAATTTGCAGCTCGGACAGCCAGCTGAAACTCTCTGGCCATTGGAGGCT 124
|||||
QY 399 GGAACCTCGGGACTGCCCTTTTCACTGCCATGGTTTCTACTCCTTGAGAACCCGACAGATT 458
|||||
Db 125 GGAACCTCGGGACTGCCCTTTTCACTGCCATGGTTTCTACTCCTTGAGAACCCGACAGATT 184
|||||
QY 459 TCATCACCCTCAGTCATCAATTCCTCGCCGAGTATGAGTTTGACGGGCTGGACTTTGACT 518
|||||
Db 185 TCATCACCCTCAGTCATCAATTCCTCGCCGAGTATGAGTTTGACGGGCTGGACTTTGACT 244
|||||
QY 519 GGGAGTACCCTCGCTCTCGTGGAGCCCTCCTCAGGACAAGCATCTTTCACCTGTCTCG 578
|||||
Db 245 GGGAGTACCCTCGCTCTCGTGGAGCCCTCCTCAGGACAAGCATCTTTCACCTGTCTCG 304
|||||
QY 579 TGCAGGAATGGCTGAAGCTTTTGGACAGAGGCCAAGCAGATCAACAAGCCCGAGGCTGA 638
|||||
Db 305 TGCAGGAATGGCTGAAGCTTTTGGACAGAGGCCAAGCAGATCAACAAGCCCGAGGCTGA 364
|||||
QY 639 TGGTCACTGCTGCTGAGTGTGCTGGCATCTCCAATATCAGCTGTGCTATGAGATCCCC 698
|||||
Db 365 TGGTCACTGCTGCTGAGTGTGCTGGCATCTCCAATATCAGCTGTGCTATGAGATCCCC 424
```

```
QY 699 AACTGTACAGTACCTGGACTACATCCATGTGTACGACCTACGACCTTCCATGGCTCTGGG 758
|||||
Db 475 AACTGTACAGTACCTGGACTACATCCATGTGTACGACCTACGACCTTCCATGGCTCTGGG 484

QY 759 AGGGGTACACTGGAGAGAACGCCCTCTTACAAATACCCGACTGACACCCGCGAGCAAG 818
|||||
Db 485 AGGGGTACACTGGAGAGAACGCCCTCTTACAAATACCCGACTGACACCCGCGAGCAAG 544

QY 819 CCTACTCAATGTGATTATGATGACTACTGGAAGGACAAATGGAGCACCAGCTGAGA 878
|||||
Db 545 CCTACTCAATGTGATTATGATGACTACTGGAAGGACAAATGGAGCACCAGCTGAGA 604

QY 879 AGCTCATCGTTGGATTCCCTACTATGGACACAACCTTCTACTCTGAGCAACCCCTCCAACA 938
|||||
Db 605 AGCTCATCGTTGGATTCCCTACTATGGACACAACCTTCTACTCTGAGCAACCCCTCCAACA 664

QY 939 CTGGAAATGGTGCCGCCCACTCTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 998
|||||
Db 665 CTGGAAATGGTGCCGCCCACTCTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 724

QY 999 GGATCTGGGCTTACTACGAGATCTGTACCTTCTGAAATAATGGAGCACTCAGGGATGG 1058
|||||
Db 725 GGATCTGGGCTTACTACGAGATCTGTACCTTCTGAAATAATGGAGCACTCAGGGATGG 784

QY 1059 ATGCCCTCTCAGGAAGTGCCTTATGCTATCAGGGCAATGTGTGGTTGGCTATGACAACA 1118
|||||
Db 785 ATGCCCTCTCAGGAAGTGCCTTATGCTATCAGGGCAATGTGTGGTTGGCTATGACAACA 844

QY 1119 TCAAGAGCTTCGATATTAAGGCTCAATGGCTTAAGCACAAATAATTTGGAGGGCCATGG 1178
|||||
Db 845 TCAAGAGCTTCGATATTAAGGCTCAATGGCTTAAGCACAAATAATTTGGAGGGCCATGG 904

QY 1179 TCTGGGCCATTCATCTGGATGACTTCACCTGGGCACTTTCTGCAACAGGCAAGTTTCCCC 1238
|||||
Db 905 TCTGGGCCATTCATCTGGATGACTTCACCTGGGCACTTTCTGCAACAGGCAAGTTTCCCC 964

QY 1239 TAATCTCCACCCTGAAGAAGGCCCTCGGCTCGAGAGTGCAGAGTTCAGAGTTCACGGCTC 1298
|||||
Db 965 TAATCTCCACCCTGAAGAAGGCCCTCGGCTCGAGAGTGCAGAGTTCAGAGTTCACGGCTC 1024

QY 1299 AGCCATTCAGCCAAATAACTGTCTGCCCTCCAGTGGGAGCGGAGCGGAGGAGTAGCA 1358
|||||
Db 1025 AGCCATTCAGCCAAATAACTGTCTGCCCTCCAGTGGGAGCGGAGCGGAGGAGTAGCA 1084

QY 1359 GCTCTGGAGGAGCTCGGGAGGAGGAGTTCCTGCTGTCTGAGAGCAACGGGCTCTACC 1418
|||||
Db 1085 GCTCTGGAGGAGCTCGGGAGGAGGAGTTCCTGCTGTCTGAGAGCAACGGGCTCTACC 1144

QY 1419 CGCTGGCAAAATACAGAAATGCTTCTGGCACTGCGTGAATGGAGTCACTACACGAGAGA 1478
|||||
Db 1145 CGCTGGCAAAATACAGAAATGCTTCTGGCACTGCGTGAATGGAGTCACTACACGAGAGA 1204

QY 1479 ACTGCCAGCGCGGCTGTGCTTCGACACAGCTGTGATGTGCTGCAACTGGGCATAAACC 1538
|||||
Db 1205 ACTGCCAGCGCGGCTGTGCTTCGACACAGCTGTGATGTGCTGCAACTGGGCATAAACC 1364

QY 1539 GACCTGGTCTATATTCCTTAGAGTTCAGTCTCTTTTGTAGGACATGTGGCCCTACC 1598
|||||
Db 1265 GACCTGGTCTATATTCCTTAGAGTTCAGTCTCTTTTGTAGGACATGTGGCCCTACC 1324

QY 1599 TAAAGTCTTGCATAAATAATCAGCAGTC 1625
|||||
Db 1325 TAAAGTCTTGCATAAATAATCAGCAGTC 1351

RESULT 3
AB025008
LOCUS Homo sapiens TSA1902-L mRNA for novel member of chitinase family,
DEFINITION complete cds.
ACCESSION AB025008
VERSION AB025008.1 GI:6467176
```


Query Match 64.4%; Score 1046; DB 9; Length 4250;

TITLE	Direct Submission
JOURNAL	Submitted (24-JUL-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK COMMENT	NIH-MGC Project URL: http://mgc.nci.nih.gov Contact: MGC help desk Email: cgabps-re@mail.nih.gov Tissue Procurement: Jeffrey E. Green, M.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center Center code: BCM-HGSC Web site: http://www.hgsc.bcm.tmc.edu/cdna/ Contact: amge@bcm.tmc.edu Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M., Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M., Richards, S., Gibbs, R.A.
	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLML at: http://image.llnl.gov Series: IRAK Plate: 24 Row: h Column: 16 This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein.
FEATURES	Location/Qualifiers
source	1..1557 /organism="Mus musculus" /db_xref="taxon:10090" /map="FVB/N" /clone="MGC:19045 IMAGE:4189080" /tissue_type="salivary gland, 10 week old female mouse" /clone_lib="NCI CGAP_SG2" /lab_host="DH10B" /note="Vector: pCMV-Sport6" 4..1425 /codon_start=1 /product="chitinase, acidic" /protein_id="AAH34548.1" /db_xref="GI:21961191" /db_xref="locusID:81600" /translation="NAKLILVTGLALLNLAQLGSAYNLICYFTWMAQYRPLGSRFPD DNPLCLTHLYAFAGMNNETITWENDVTLRYAFNDLKNRNLSKLATLLAIGWNFG TAPEFTWSTNSORFTFTSVIKFLROYFDGLDLWEYPGSRSPQDKHLFTLVVK EMREAFOEAIENSPRLMTVAAYAGGISNQAGEYEIPELSKYLDFHVHTYDLHGSH EGYTGENSLPKYIPETSGSNAYLNVDYVNTWKNGAPAEKLIYGFEPHYHTFILRP SDNGIAQDGDPAGPYTRQAGWAYEIEICFLRSRGATEVMDASQBPVYKANEWL GYDNIKPSVKAQKLNQNFQGMIAWIDLDDFTFSFGDQKGKPLPILTPYALKAIGISTE GTADVPVSEPVTPTPPGGSGGSSGSSGSGSCADKADGLYPVADRANFWQING ITYOQHCOAGLVEDTSCNCCNP"
CDS	BASE COUNT 410 a 382 c 366 g 399 t ORIGIN
	Query Match 64.3%; Score 1045.6; DB 10; Length 1557; Best Local Similarity 83.5%; Pred. No.: 3e-270; Matches 1202; Conservative 0; Mismatches 229; Indels 9; Gaps 1;
QY	101 ACCATGACAAAGCTATTCTCCTCACAGTCCTTGCTCTTTACTGCAATTCACAGCTCGGC 160
DB	1 ACAGTGGCAAGCTACTTCTCGCTACAGCTCTGGCTCTTCTGCTGAATGCTCAGCTGGG 60
QY	161 TCTGCTACCAGCTGACATGCTTACCACTGGGCCAGTAGTACCGGCCAGGCTGGG 220
DB	61 TCTGCTTACAATCTGATATGCTATTTCACCACTGGGCCAGTAGTACCGGCCAGGCTGGG 120
QY	221 CGCTTCATCCCTGACAACTACACCCCTCGCTCTGTACCCACCTGATCTAGCGCTTTGCT 280
DB	121 AGCTTCAAGCCTGATGACATAAACCCCTGCCTGTGTACTCACTGATCTATGCCTTTGCT 180
QY	281 GGAGGCGAGAACAACAGATACACCACTACGAATGGAACGATGTGACTCTTACCAAGCT 340
DB	181 GGATCGACAACAANTGATACCACTACCACTAGATGAATGATGTACTCTTATTAAGCT 240


```
Qy 948 GTGCCCCACCTCTGGTGTCTCTCTGGGCGCTATGCCAAGAGCTCTGGGATCTGGG 1007
Db 842 GTGCCCCACCTCTGGTGTCTCTGGGCGCTATGCCAAGAGCTCTGGGATCTGGG 901
Qy 1008 CTTACTACGAGACTGTACCTTCTTGAAATGCGAGCCACTCAGGAGTGGATGCCCTC 1067
Db 902 CCTACTAGAGATTGACCTTCTTGAGAGTGGAGCCACTGAGGCTGGGATGCCCTCC 961
Qy 1068 AGAAGTGCCTTATGCTATCAGGCGCAATGTGTGGTGGCTATGACAACATCAAGAGCT 1127
Db 962 AGAAGTGCCTTATGCTATCAGGCGCAATGTGTGGTGGCTATGACAACATCAAGAGCT 1021
Qy 1128 TCGATATTAAGGCTCAATGCTTAAAGCAACAATAATTGGAGCGCCATGCTCTGGGCCA 1187
Db 1022 TCACTGTAAAGGCTCACTGCTTAAAGCAGACAATAATTGGAGGTGCCATGATCTGGGCCA 1081
Qy 1188 TTGATCTGGATGACTTCACTGGCACCTTCTGCAACAGGCGCAAGTTTCCCTATCTCA 1247
Db 1082 TTGACCTGTGATGACTTCACTGGCTCTTCTGTGATCAGGGAATAATTCTCTGACTTCTA 1141
Qy 1248 CCCTGAAGAAGCGCTCGGCTGCGAGAGTGCAGAGTGCAGGCTCCAGCTCAGCCCATG 1307
Db 1142 CTTTGACAAGCCCTTGGCATATCCACTGAAGTTCGACAGCTCTGAGCTCGCTCCG 1201
Qy 1308 AGCAATAACTGTCTCCAGTGGCAGCGGAGCGGAGCGGAGTAGCAGCTCTGGAG 1367
Db 1202 AGCAGTGACTACTCTCC-----AGGAAGTGGGAGTGGGGTGAAGCTCCGGAG 1252
Qy 1368 GCAGCTCGGAGGAGTGGATTCTGTCTGTCAGAGCAACGCGCTTACCCCGTGCCAA 1427
Db 1252 GAAGCTCTGGAGGAGTGGATTCTGTCCGACAAAGCAGATGGCCTTACCCCTGTGGCAG 1312
Qy 1428 ATACAGAAATGCTTCTGGCACTGCGTGAATGGAGTCACTACGACAGAACTGCCAGG 1487
Db 1312 ATGACAGAAATGCTTCTGGCACTGCGTGAATGGAGTCACTACGACAGAACTGCCAGG 1427
Qy 1488 CCGGCTGTCTTGACACAGCTGTGATTGCTGCAACTGGGCATAAACCTGA 1540
Db 1373 CAGGCTGTCTTGATACACAGCTGTGATTGCTGCAACTGGGCATAAACCTAA 1425

RESULT 9
AF290003 LOCUS 1530 bp mRNA linear ROD 27-FEB-2001
DEFINITION Mus musculus acidic mammalian chitinase precursor, mRNA, complete
cgs.
ACCESSION AF290003
VERSION AF290003.1 GI:12597290
KEYWORDS
SOURCE J. Biol. Chem. 276 (9), 6770-6778 (2001)
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1530)
Boot,R.G., Blommar, E.F., Swart, E., Ghaubari-van der Vlugt, K.,
Blj, N., Moe, C., Place, A. and Aerts, J.M.
Identification of a novel acidic mammalian chitinase distinct from
chitotriosidase
J. Biol. Chem. 276 (9), 6770-6778 (2001)
MEDLINE 21125893
PUBMED 11085997
REFERENCE 2 (bases 1 to 1530)
AUTHORS Boot, R.G., Verhoeck, M., Swart, E. and Aerts, J.M.
TITLE Direct Submission
JOURNAL Submitted (26-JUL-2000) Dept. of Biochemistry, Academic Medical
Center, University of Amsterdam, Meibergdreef 15, Amsterdam 1105
AZ, The Netherlands
FEATURES
source Location/Qualifiers
1.1530
/organism="Mus musculus"
/strain="BALB/c"
/db_xref="taxon:10090"
/sex="male"
```

CDS

```
/tissue_type="lung"
/dev_stage="9-11 weeks"
3.1424
/EC_number="3.2.1.14"
/note="chitinase precursor; AMCase"
/codon_start=1
/product="acidic mammalian chitinase precursor"
/protein_id="AAG60018.1"
/db_xref="GI:12597291"
/translation="MAKLLVLTGLALLNLAOLGSAYNLICFTNWAQYRPLGSEFKPD
DINPCLTHLIYAFAGMONEITIEDNDVTLKAFNDLKNRNSKLTLLAIGWNFG
TAPFTMTVSTSONRQFTITSVKELROYGDFDGLDWEYFPGSRGPPQDKHLEFVLVW
EMREAFEEAIESNPRMLVTAAGVINSIQAGYEIPESLYLDTIHVNTYDLHGSV
EGYTGNSPLKYPTETGSNAYLVNDVVMYKNKNGAPAEKLVGPPEYFKLRNP
SDNGIGAPTSDDGPAGATYRQAGFWAYEICTFLRSGATEVWDASOEVYAYKANWL
GYDNIKSFYKAWLKNONFEGAMWAILDDLDGSCDQGFPLSTLNKALGISTE
GCTAPDVPSEPTTPPGSGSGSGSGSGSGSGGFCADKADGLIPVADDRNAPFOCING
ITYQQHQAGLGVDFDTSNCCNWP"
BASE COUNT 382 a 383 c 366 g 399 t
ORIGIN
Query Match 64.1%; Score 1041.4; DB 10; Length 1530;
Best Local Similarity 83.3%; Pred. No. 4.1e-269;
Matches 1199; Conservative 0; Mismatches 231; Indels 9; Gaps 1;
Qy 102 CCATGACAAAGCTTATCTCTCACAGGCTCTTGCTCTTATCTATGAAATTCAGCTCGCT 161
Db 1 CGATGCCAAGCTACTTCTCTCACAGGCTCGCTCTCTCTCTGAAATGCTCAGCTGGGT 60
Qy 162 CTGCTACACAGCTGACATGCTACTTACCAACTGGGCCCCAGTACCAGGCTGGGGC 221
Db 61 CTGCTACAATCTGATGCTATTTTACCAACTGGGCCCCAGTATCGGCCAGGTCTGGGGA 120
Qy 222 GCTTCATGCTGACACATCGACCCCTGCTCTGTACCCACCTGATCTAGCCTTTGCTG 281
Db 121 GCTTCAGGCTGATGACATTAACCCCTGCTGTACTACCTGATCTATGCTCTTCTG 180
Qy 282 GGAGGAGAGAACACGAGATCACCACCATCGAATCGAATGGAACGATGTGACTCTTACCAAGCTT 341
Db 181 GGATGAGAACATGATGATCACCACCATAGATGGAATGGAATGATGTTACTCTCTATAAAGCTT 240
Qy 342 TCAATGGCTGAAATAAGAACAGCAGCTGAAACTCTCTGGGCCATTTGGAGGCTGGA 401
Db 241 TCAATGACTTGAATAAGAACAGCAGCAAACTGAAACCTCTCTGGGCAATTTGGAGGCTGGA 300
Qy 402 ACTTGGGAGCTGCCCTTTTCCATGCCATGCTTCTACTCTCTGAGAACCGCAGACTTTCA 461
Db 301 ACTTGGAGCTGCTCTTCTCCTACCATGCTTTTCCACTTCTCAGAACCGCCAGACTTCA 360
Qy 462 TCACCTCAGTCATCAAAATTCCTGCGCCAGTATGAGTTTGACGGGCTGGACTTTGACTGGG 521
Db 361 TTACCTCAGTCATCAAAATTTCTCGTCAGTATGGTTTGTATGACTGGACTGGAGTGG 420
Qy 522 AGTACCTGCTCTCTGTTGGAGCCCTCTCTCAGAACAGCATCTCTTCAGTCTGCTGGTGC 581
Db 421 AATACCCAGGCTCACGCTGGGAGCCCTCTCAGAACAGCATCTCTTCAGTCTGCTGGTGA 480
Qy 582 AGGAAATGCTGAGGCTTTTGGAGCAGGAGCCAGCAGATCAACAAGCCAGCCTGATGG 641
Db 481 AGGAAATGCTGAGGCTTTTGGAGCAGGAGCTATTGAGACCAACAGCCAGCATGATGG 540
Qy 642 TCACCTGCTGAGTATGCTGCTGGCATCTTCCAAATATCCAGTCTGGCTATGAGATCCCCAAC 701
Db 541 TTACTGCTGCTGATGCTGGGATTTTCCAAATATCCAGGCTGGCTATGAGATCCCCGAAC 600
Qy 702 TGTACAGTACCTGGAGTACATCATGATGATGATGATGATGATGATGATGATGATGATGATG 761
Db 601 TTTCTAAGTACCTGGATTTTCTCATGATGATGATGATGATGATGATGATGATGATGATGATG 660
Qy 762 GCTTACACTGGAGAGACAGCCCTCTTACAAATATCCAGTCTGACACCGGAGCAGCAAGCCT 821
Db 661 GCTTACACTGGGAGAGATGCTCTTTACAAATATCCCTTACTGAGACTGGTATGAGCAATGCCT 720
```

```

822 ACCTCAATGGGATTTATGTCATGAACCTACTGGAAGGACAATGAGCACCAGCTGAGAAGC 881
Db      |||
721 ACCTCAATGGGATTTATGTCATGAACCTACTGGAAGGACAATGAGCACCAGCTGAGAAGC 780
QY      |||
882 TCATCGTTGGATTCCTACCTATGAGCACAACCTTATCTGAGCACAACCCCTCCACACTG 941
Db      |||
781 TCATCGTTGGATTCCTACCTATGAGCACAACCTTATCTGAGCACAACCCCTCCACACTG 840
QY      |||
942 GAATTTGGTGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1001
Db      |||
841 GAATTTGGTGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
QY      |||
1002 TCTGGCTTACTACGAGATCTGATCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1061
Db      |||
901 TCTGGCTTACTACGAGATCTGATCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
QY      |||
1062 CCCCTCAGGAAGTGGCTTATGCTATCAGGGAATGTTGGTGGTGGTGGTGGTGGTGGT 1121
Db      |||
961 CCTCCCAAGAAGTGGCTTATGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
QY      |||
1122 AGAGCTTCGATATTAGGCTCAATGCTTAAAGCACAACAAATTTGGAGCGCCATGGTCT 1181
Db      |||
1021 AGAGCTTCGATATTAGGCTCAATGCTTAAAGCACAACAAATTTGGAGCGCCATGGTCT 1080
QY      |||
1182 GGGCCATTGATGCTGATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1241
Db      |||
1081 GGGCCATTGATGCTGATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
QY      |||
1242 TCTCCACCTGAAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1301
Db      |||
1141 CTTCTACTTTGAACAAGCCCTTGGCATATCCACTGAAGGTTGCACAGCTCCTGACGTGC 1200
QY      |||
1302 CCATTGAGCAATAACTGCTGCCAGTGGCAGCGGGAACGCGGAGCGGAGCTACAGCT 1361
Db      |||
1201 CTTCCGAGCGACTACTCTCC-----AGGAGTGGGAGTGGGGTGGAGCT 1251
QY      |||
1362 CTGGAGGAGCTCGGGAGGAGTGGATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1421
Db      |||
1252 CCGGAGGAGCTCTGGAGGAGTGGATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1311
QY      |||
1422 TGGCAAAATACAAATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1481
Db      |||
1312 TGGCAGATGACAGAATGCTTTTGGCAGTGGCATCAATGGAATCATACACAGCAGCAT 1371
QY      |||
1482 GCGAGCGCGGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1540
Db      |||
1372 GTCACAGCAGGCTGCTTTTGTATACAGCTGTAATGCTGCAACTGGCCATGAACCTAA 1430

RESULT 10
AB025009
LOCUS   Homo sapiens TSA1902-S mRNA for novel member of chitinase family,
DEFINITION complete cds.
ACCESSION AB025009
VERSION   AB025009.1 GI:6467178
KEYWORDS novel member of chitinase family; TSA1902-S.
SOURCE    Homo sapiens cDNA to mRNA.
ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (sites)
AUTHORS   Saito,A., Ozaki,K., Fujiwara,T., Nakamura,Y. and Tanigami,A.
TITLE     Isolation and mapping of a human lung-specific gene, TSA1902,
           encoding a novel chitinase family member
JOURNAL   Gene 239 (2), 325-331 (1999)
MEDLINE   20018184
REFERENCE 2 (bases 1 to 1188)
AUTHORS   Saito,A., Ozaki,K., Fujiwara,T., Takahashi,E. and Tanigami,A.
TITLE     Direct Submission
JOURNAL   Submitted (16-MAR-1999) Akihiko Saito, Otsuka Pharmaceutical Co.,
           Ltd., Otsuka GEN Research Institute; 463-10 Kagasuno Kawauchi-cho,
           Tokushima, Tokushima 771-0192, Japan (E-mail:saito@otsuka.gr.jp,
```

```

Tel:81-88-665-2888, Fax:81-88-637-1035)
FEATURES
Source    Location/Qualifiers
           1..1188
           /organism="Homo sapiens"
           /db_xref="taxon:9606"
           1..1188
           /gene="TSA1902-S"
           150..1097
           /gene="TSA1902-S"
           /codon_start=1
           /product="novel member of chitinase family"
           /protein_id="BAA86981.1"
           /db_xref="GI:6467178"
           /translation="MREAFQEAQKQINPRLMVTAAVAGISNIOGVEIPOLSOYLD
           FYHVTYDLHGSWEGYTGENSEPLKYPTDGSNAYLNDVYVMYKWDNCAPAEKLIIV
           PTHGNFILSNPSTIGAPTSAGPAGPYKESGIWAYIEICITFLKNGATQGWDAIP
           QESTPAYGNGVWGYNDYKSFIDKAOMLKHKFGAMWAIIDLDFDTFTFCNOGKPL
           IETLKLALGLASASCTAPQPIETPTAAPSNGSGSGSSSSSGSGSGSGGFCAGRANGL
           YPVANNRANFHVHCVNGVYQQNCQAGLVFDTSCDCNNA"
BASE COUNT      280 a      337 c      309 g      262 t
ORIGIN
Query Match      63.8%; Score 1037.2; DB 9; Length 1188;
Best Local Similarity 99.7%; Pred. No. 5.4e-268;
Matches 1039; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY      584 GAAATGCGTGAAGCTTTTGAGCAGGAGGCCAAGCAGATCAACAAGCCAGGCTGATGGTC 643
Db      |||
147 GAAATGCGTGAAGCTTTTGAGCAGGAGGCCAAGCAGATCAACAAGCCAGGCTGATGGTC 206
QY      644 ACTGCTGCAGTAGCTGCTGGCATCTCCAATATCCAGTCTGGCTATGAGATCCCCAACTG 703
Db      |||
207 ACTGCTGCAGTAGCTGCTGGCATCTCCAATATCCAGTCTGGCTATGAGATCCCCAACTG 266
QY      704 TCACAGTACCTGGATACATCATGTACGACTACGACCTCCATGGCTCCCTGGAGGGC 763
Db      |||
267 TCACAGTACCTGGATACATCATGTACGACTACGACCTCCATGGCTCCCTGGAGGGC 326
QY      764 TACACTGGAGAGAACAGCCCCCTCTACAAATACCCGACTGACACCGGAGCAGCCTAC 823
Db      |||
327 TACACTGGAGAGAACAGCCCCCTCTACAAATACCCGACTGACACCGGAGCAGCCTAC 386
QY      824 CTCATATGGGATTATGTCATGAATCTACTGGAAGGACAATGGAGCACCAGCTGAGAAGCTC 883
Db      |||
387 CTCATATGGGATTATGTCATGAATCTACTGGAAGGACAATGGAGCACCAGCTGAGAAGCTC 446
QY      884 ATCGTTGGATTCCTTACCTATGGACACAACTTCACTCTGAGCAACCCCTCCCAACTGGA 943
Db      |||
447 ATCGTTGGATTCCTTACCTATGGACACAACTTCACTCTGAGCAACCCCTCCCAACTGGA 506
QY      944 ATTGTTGCCCCCACCCTCTGGTCTGCTGGGCCCCATGTCGAAGGAGTCTGGGATC 1003
Db      |||
507 ATTGTTGCCCCCACCCTCTGGTCTGCTGGGCCCCATGTCGAAGGAGTCTGGGATC 566
QY      1004 TGGGCTTACTACGAGATCTGTACCTTCTGAAAAATGGAGCCACTCAGGAGTGGGATGCC 1063
Db      |||
567 TGGGCTTACTACGAGATCTGTACCTTCTGAAAAATGGAGCCACTCAGGAGTGGGATGCC 626
QY      1064 CTTCAAGAGTGGCTTTATGCTTATCAGGCAATGTGTGGTGGTGGTGGTGGTGGTGGT 1123
Db      |||
627 CTTCAAGAGTGGCTTTATGCTTATCAGGCAATGTGTGGTGGTGGTGGTGGTGGTGGT 686
QY      1124 AGCTTCGATATTAAAGCTCAATGGCTTAAGCACAACAAATTTGGAGGCGCCATGGTCTGG 1183
Db      |||
687 AGCTTCGATATTAAAGCTCAATGGCTTAAGCACAACAAATTTGGAGGCGCCATGGTCTGG 746
QY      1184 GCATTGATCTGGATGACTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1243
Db      |||
747 GCATTGATCTGGATGACTTCACTGGCCTTCTGCAACAGGCGCAAGTTTCCCTTAATC 806
QY      1244 TCCACCTCGAAGAGCGCCCTCGGCTGCGAGTGCAGAGTGCAGGCTCCAGCTCAGGCC 1303
Db      |||
807 TCCACCTCGAAGAGCGCCCTTGGCTTGGCCTTGCAGAGTGCAGAGTGCAGGCTCCAGCTCAGGCC 866
```

```
QY 1304 ATTGAGCAATAACTGCTGCCAGTGGCAGCGGGAACGGAGCGGAGTAGCAGCTCT 1363
|||||
Db 867 ATTGAGCAATAACTGCTGCCAGTGGCAGCGGGAACGGAGCGGAGTAGCAGCTCT 926
|||||
QY 1364 GGAGCGAGCTGGGAGGAGTGGATTCTGCTGTGACAGCGCAACGGGCTCTACCCCGTG 1423
|||||
Db 927 GGAGCGAGCTGGGAGGAGTGGATTCTGCTGTGACAGCGCAACGGGCTCTACCCCGTG 986
|||||
QY 1424 GCAATAACAGAAATGCTTCTGCGACTGCGTGAATGGAGTCACTACCGAGCAAGTGC 1483
|||||
Db 987 GCAATAACAGAAATGCTTCTGCGACTGCGTGAATGGAGTCACTACCGAGCAAGTGC 1046
|||||
QY 1484 CAGCCCGGGCTGTGCTGACACACGAGTGTGATTGCTGCACTGGGCAATACCTGACCT 1543
|||||
Db 1047 CAGCCCGGGCTGTGCTGACACACGAGTGTGATTGCTGCACTGGGCAATACCTGACCT 1106
|||||
QY 1544 GGCTATATATCCCTAGAGTCCAGTCTCTTTGCTTAGGACATGTTGCCCTACTAAAG 1603
|||||
Db 1107 GGCTATATATCCCTAGAGTCCAGTCTCTTTGCTTAGGACATGTTGCCCTACTAAAG 1166
|||||
QY 1604 TCCTGCAATAAAATCAGCAGTC 1625
|||||
Db 1167 TCCTGCAATAAAATCAGCAGTC 1188
|||||

RESULT 11
LOCUS M94584 1525 bp mRNA linear ROD 15-JUN-2001
DEFINITION Mus musculus secretory protein precursor (Yml) mRNA, complete cds.
ACCESSION M94584
VERSION M94584.2 GI:11140876
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1525)
Chang,N.-C.A., Liu,C.-H. and Chang,A.C.
Molecular characterization of a secretory protein (Ym-1)
transiently expressed by activated murine peritoneal macrophages
Unpublished
2 (bases 1 to 1525)
Chang,N.C., Hung,S.I., Hwa,K.Y., Kato,I., Chen,J.E., Liu,C.H. and
Chang,A.C.
A macrophage protein, Yml, transiently expressed during
inflammation is a novel mammalian lectin
J. Biol. Chem. 276 (20), 17497-17506 (2001)
21264517
11297523
3 (bases 1 to 1525)
Chang,N.-C.A., Liu,C.-H. and Chang,A.C.
Direct Submission
Submitted (27-APR-1993) Microbiology and Immunology, National
Yang-Ming University, 155, Sect. II, Li-Noon St, Taipei, Taiwan
112, R.O.C.
4 (bases 1 to 1525)
Chang,N.-C.A., Liu,C.-H. and Chang,A.C.
Direct Submission
Submitted (13-NOV-2000) Microbiology and Immunology, National
Yang-Ming University, 155, Sect. II, Li-Noon St, Taipei, Taiwan
112, R.O.C.
REMARK
Sequence update by submitter
COMMENT On Nov 13, 2000 this sequence version replaced gi:202441.
FEATURES
source 1..1525
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone_lib="PEC in lambda gt11 and ZAPII"
/dev_stage="adult"
1..1525
/gene="Yml"
13..1209

QY 101 ACATGACAAAGCTTATCTCCACAGGCTGCTCTTATCTATCTGATGATGAGCTCGGC 160
|||||
Db 10 ACATGACAAAGCTTATCTCTGACAGGCTGCTCTTATCTGATGATGAGCTCGGC 69
|||||
QY 161 TCTGCTTACCAGCTGACATGCTACTTCCACCACTGGCCGAGTACCGCCGAGGCTCGGG 220
|||||
Db 70 TCTTCTTACCAGCTGATGCTGCTACTATACCACTGGGCTGAGGACAGGCAATAAGAGGG 129
|||||
QY 221 CGCTTCTGCTGACCAATCGACCCCTGCTCTGTACCCACCTGATCTACGCTTTGCT 280
|||||
Db 130 AGTTTCAACCTGCTGATATGACCCCTGCTGCTGCTACTCACCTGATCTATGCTTTGCT 189
|||||
QY 281 GGAGGAGCAGAACAGGAGATCACCACCATCGAATGGAAACGATGTGATCTCTACCAAGCT 340
|||||
Db 190 GGAATGCAAGTAATGAGATCACTTACACATGAGCAAGACTTGCCTGACTATGAAGCA 249
|||||
QY 341 TTCAATGGCTGAAATAAAGACAGCAGCTGAAACTCTCTGGCCATTGGAGGCTGG 400
|||||
Db 250 TTGATGGCTGAAAGACAGACACTGAGCTAAAACTCTCTGGCCATTGGAGGATGG 309
|||||
QY 401 AACTTGGGAGTGCCTTTTCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 460
|||||
Db 310 AAGTTTGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 369
|||||
QY 461 ATCACTCAGTCACTCAAAATCTCTGCGCCAGTATGAGTTTACGGGCTGAGCTTTGACTGG 520
|||||
Db 370 ATTCACTCAGTCACTCAAAATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 429
|||||
QY 521 GAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 580
|||||
Db 430 CAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 489
|||||
QY 581 CAGGAATGCTGAGGCTTTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 640
|||||
Db 490 AAGGAATGCTGAGGCTTTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 549
|||||
QY 641 GTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 700
|||||
Db 550 CTCACTTCCACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 609
|||||
QY 701 CTGTCACAGTACCTGAGTACCTGAGTACCTGAGTACCTGAGTACCTGAGTACCTGAGTACCTGAG 760
|||||
Db 610 CTGTCACAGTACCTGAGTACCTGAGTACCTGAGTACCTGAGTACCTGAGTACCTGAGTACCTGAG 669
|||||
```

```
QY 761 GGCTACACTGGAGAACAGCCCTCTACAAATACCGACAGCCGACAGCC 820
Db 670 GGCTACACTGGAGAAATAGTCCCTCTATAATATCCATATGGAAGAGTGT 729
QY 821 TACCTCAATGTGGATATGCTGAATCACTGAGGAGCAATGAGCACCAGCTGAGAAG 880
Db 730 GATCTCAATGTGGATATGCTGAATCACTGAGGAGCAATGAGCACCAGCTGAGAAG 789
QY 881 CTCATCGTGGATTCCTACCTATGAGACAACTCATCTGAGCAACCCCTCCACATC 940
Db 790 CTCATCGTGGATTCCTACCTATGAGCATACTGAGCATACTCTGAGCACCCTCTAGACT 849
QY 941 GGAATGPGCCCCCACCCTCTGCTGCTGCTGCGGCCCTATGCCAAGGAGTGGG 1000
Db 850 GGAATGPGTCCCTACAATAGTACTGCCCCCACCAGGAAAGTACACAGATGAATCAGCA 909
QY 1001 ATCTGGGCTTACTACGAGATCTGATCTTCTGTAATATGAGGAGGAGGAGGAGTGGAT 1060
Db 910 CTCCTGGCTTACTATGAGTGTGATCAATTTCTACTGGAAGGAGGAGGAGGAGTGGAT 969
QY 1061 GCGCCCTCAGGAAGTGCCTTATGCTATCAGGCAATGTGTGGTGGCTATGACAACTC 1120
Db 970 GCGCCCCAGGAGTACCTTATGCTATCAGGTAATGAGTGGTGGTATGACATGTC 1029
QY 1121 AAGAGCTTCGATATTAAGCTCAATGGCTTAAGCACAACAAATTTGAGGCGCCATGGTC 1180
Db 1030 AGGAGCTTCAAGTTGAAGCTCAGTGGCTCAAGGACAACAAATTTAGGAGGTGCGGTGTC 1089
QY 1181 TGGGCAATGATCTGATGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTG 1240
Db 1090 TGGGCCCCGATGATGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTG 1149
QY 1241 ATCTCCACCCTGAAGAGGCCCTCGGCTGAGAGTCAAGTTGACGCGCTC 1293
Db 1150 ACATCTACTTAAAGGAGATCTCAATATACAGTGAAGTTGCAAGGCC 1202
```

RESULT 12

```
D87757 1506 bp mRNA linear ROD 12-JAN-2000
Mus musculus mRNA for ECF-L precursor, complete cds.
D87757
ECF-L precursor.
Mus musculus bone marrow cdna to mRNA.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Ohashi.M., Arita.H. and Hayai.N.
Identification of a novel eosinophil chemotactic cytokine (ECF-L)
as a chitinase family protein
J. Biol. Chem. 275 (2), 1279-1286 (2000)
20092902
Ohashi.M.
Direct Submission
Submitted (06-SEP-1996) Makoto Ohashi, The University of
Tokushima, Faculty of Integrated Arts and Sciences; 1-1
Minami-Johsanjima, Tokushima, Tokushima 770, Japan
(E-mail:ohashieias.tokushima-u.ac.jp, Tel:0886-56-7261,
Fax:0886-56-7298)
On Oct 7, 1999 this sequence version replaced gi:1545818.
Sequence updated (05-Oct-1999).
FEATURES
source
1. .1506
/organism="Mus musculus"
/db_xref="taxon:10090"
/tissue_type="bone marrow"
1. .1506
/gene="mECF-L"
1. .1197
/gene="mECF-L"
```

Db 781 ATTCTGGGATTTCCAGCATATATGGGCAATACCTTTATCTAGTACCCCTTTCAAGACTGGA 840

Qy 944 ATTGGTCCCCCACTCTGGTCTGGTCTGGTCTGGGCCCCTATGCAAGGAGCTGGGATC 1003

Db 841 ATTGGTCCCCCACTCTGGTCTGGTCTGGTCTGGGCCCCTATGCAAGGAGCTGGGATC 900

Qy 1004 TGGGCTTACTACGAGATCTGTACCTCTCTGAAATAGGACCTACGAGGATGGATGCC 1063

Db 901 CTGGCTTACTATGAGTCTGTACATTTCTGAATGAAGGAGCCACTGAGGTCTGGATGCC 960

Qy 1064 CCTCAGGAAGTCCCTATGCTATCAGGCAATCTGGTGGTGGCTATGACCAATCAAG 1123

Db 961 CCCGGAAGTACCTATGCTATCAGGGAATATAGTGGTGGTGGTATGACCAATGTCAGG 1020

Qy 1124 AGCTTCGATATTAAGGCTTAAGGCTTAAGCAACAAATTTGGAGGCGCATGGTCTGG 1183

Db 1021 AGCTTCAGTGAAGGCTCAGTGGCTCAAGGACAAATTTAGGAGGCGGTGGTCTGG 1080

Qy 1184 GCATTCATCTGGATGACTTCACTGGCACTTTTCGCAACAGGCGCAAGTTTCCCTCAATC 1243

Db 1081 CCGTGTGACATGGATGACTTCACTGGTCTTTCTGTCCAGAGACATTTCCCTCTGACA 1140

Qy 1244 TCCACCTGTAAGAGGCGCTCGGCTCAGAGTGCAGTTCGCAAGTTCGAGGCTCC 1293

Db 1141 TCTACTTAAGGAGGATCTCAATATACAGTGCAGTTCGCAAGTTCGAGGCGCC 1190

RESULT 13

AY049765 1209 bp mRNA linear ROD 06-AUG-2002

LOCUS Mus musculus secreted protein precursor Ym2 mRNA, complete cds.

DEFINITION AY049765

ACCESSION AY049765.2 GI:22123906

VERSION

KEYWORDS

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 1209)

AUTHORS Webb,D.C., McKenzie,A.N. and Foster,P.S.

TITLE Expression of the Ym2 lectin-binding protein is dependent on interleukin (IL)-4 and IL-13 signal transduction: identification of a novel allergy-associated protein

J. Biol. Chem. 276 (45), 41969-41976 (2001)

JOURNAL 21551268

MEDLINE 11553626

PUBMED

REFERENCE 2 (bases 1 to 1209)

AUTHORS Webb,D.C. and Foster,P.S.

TITLE Direct Submission

JOURNAL Submitted (02-AUG-2001) Biochemistry and Molecular Biology, John Curtin School of Medical Research, Australian National University, Off Mills Rd, Acton, Canberra, ACT 2601, Australia

REFERENCE 3 (bases 1 to 1209)

AUTHORS Webb,D.C. and Foster,P.S.

TITLE Direct Submission

JOURNAL Submitted (06-AUG-2002) Biochemistry and Molecular Biology, John Curtin School of Medical Research, Australian National University, Off Mills Rd, Acton, Canberra, ACT 2601, Australia

REMARK Sequence update by submitter

COMMENT On Aug 6, 2002 this sequence version replaced gi:15705156.

FEATURES

Location/Qualifiers

1..1209

/organism="Mus musculus"

/strain="BALB/c"

/db_xref="taxon:10090"

/note="isolated from the lung lavage from allergic mice"

1..1209

/note="expression regulated by IL-4 and IL-13 signaling via the IL-4R alpha subunit"

/codon_start=1

/product="secreted protein precursor Ym2"

/protein_id="AAL03953.2"

/db_xref="GI:22123907"

CDS

/translation="MAKLIILVTGLAILNLVOLGSSYQLMCIYTSWAKDRPTGSGFKPG
NIDPCLTHIYAFAGMKNEIYLSQDLRDYALNGLDRNLTLLAIGWKKFQ
PAPSSMVSTPQNRQTEIKSVIRFQYDFGLNDLMDQYPCSRGSPKDKHLFSLVQ
EMKAPFEESTLNIHPELLTSCAGIDVTKSYKTPELSQSLDYLOVMTYDLHDPK
NGTGENSPLYKSPYDICKSADLNVDSIIITYWKDHGAASEKLIYGVFPAYGHTFILSDP
SKNGIDGPTVSAGPPGYITNEQGLLAYFEICTELNEGATEIFDQTEVPYAILGNEW
GYDNVSRFKLKAQWLKDNNLGGAVVWPLMDDFSGSFHQGRPLPTTLTKRLDNVHSA
SCRASVYRGEL"

sig_peptide 326 a 273 c 282 g 328 t

BASE COUNT 1..63

ORIGIN

Query Match 42.4%; Score 689.4; DB 10; Length 1209;
Best Local Similarity 73.6%; Pred. No. 2.5e-174;
Matches 879; Conservative 0; Mismatches 316; Indels 0; Gaps 0;

Qy 104 ATGACAAAGCTTATTCCTCCACAGGCTTCTGCTTATACCTATGATTTGCGAGCTCGGCTCT 163

Db 1 ATGGCCCAAGCTCATTTCTGTACAGGCTGCGCAATCTTCTGAATGTACAGCTGGGATCT 60

Qy 164 GCCTACCAGCTGACATCTACTTACCAACTGGGCCAGTACCGCCAGGCGCTGGGCGC 223

Db 61 TCCTACCAGCTGATGTCTACTATACAGCTGGCTGAAGCAGGCGCCACAGAGGAGT 120

Qy 224 TTCTAGCTGACAAATCGACCCCTGCTCTGTACCCACCTGTACTACGCTTTTGTCTGG 283

Db 121 TTCAAACTGTATAATTGACCCCTGCTGTACTCACCTGATCTATGCTTTGCTGG 180

Qy 284 AGCAGAACACAGAGATCACCACTCAATGAGAACGATGTACTCTTACCAAGCTTC 343

Db 181 ATGAAGAATAATGAGATCACTTAAAGTGAGCAAGACTTGGTGTACTATGAAGCATTA 240

Qy 344 AATGGCTGAAAATAAGAACACAGCAGCTGAAAACCTCTCTGGCATTGGAGGCTGGAAC 403

Db 241 AATGGCTGAAAACAGCAGAACACTGAGCTAAAACCTCTCTGGCATTGGAGGATGGAAG 300

Qy 404 TTGGGACTGCCCTTTTACTGCCATGGTTTCTACTCTCTAGAACACCCAGACTTTCATC 463

Db 301 TTGGGACTGCCCTTTTACTGCCATGGTTTCTACTCTCTAGAACACCCAGACTTTCATC 360

Qy 464 ACCTCAGTCAATAAATCTCGCCAGTATGAGTTTGGAGGCTGGAGTTTGGAGGAG 523

Db 361 AAGTCAGTTATCAGATTCCTTCGTCAATATACTTTGATGGCTTCAACCTGGAGTGGCAG 420

Qy 524 TACCTGGCTCTCTGGGAGCCCTCTCTCAGAACAGCATCTTCTACTGCTCTGGTGGCAG 583

Db 421 TACCTGGGCTCTCAGGAGCCCTCTCTAAGSACAAACATCTCTTCACTGTTCTGGTGGCAG 480

Qy 584 GAAATGGTGAAGCTTTTGGAGGAGGCGCCACAGATCAACAGCCAGGCTGTAGGTC 643

Db 481 GAAATGGTGAAGCTTTTGGAGGAGGCGCCACATTCCTTGAACACATTCCTCAAGGCTGCTACT 540

Qy 644 ACTGTGCTAGCTGCTGGCCTCTCCAATATCAGTCTGGCTATGAGATCCCACTG 703

Db 541 ACTTCCAGAGCTGGATTCATTTGAGTTCAGTTCATGAGTTCGGTACAGATCCCTGAACTG 600

Qy 704 TCACAGTACCTGGACTACATCCATGTCATGACCTAGCAGCTCCATGCTCTGGAGGCG 763

Db 601 TCTCAGTCTCTCAGCTATATTCAGGTCATGACATATGATCTCCATGATCTTAAGATGGC 660

Qy 764 TACACGGAGAGAACAGCCCTCTTACAATACCCGAGTACACCGGAGCAGACCGCTAC 823

Db 661 TACACTGGAGAAAATAGTCCCTCTATAAATCTCCATATGACATTTGAAAGAGTGTGAT 720

Qy 824 CTCATGTGATTTATCTGACTGACTACTGAGGAGCAATGGAGCAGCAGCTGAGAGCTC 883

Db 721 CTCATGTGATTTCAATTTATCTACTGAGGAGCAGCTGAGGAGCTTCTTGAGAGCTC 780

Qy 884 ATGCTTGGATTTCCCTACCTATGGACACAACTTCTCATCTGAGCAACCCCTCCCACTGGA 943

Db 781 ATTGTGGGATTTCCACACATATGTCATACCTTTATCTCTGAGTACCCCTTCTAAGATGGA 840

Qy 944 ATTGGTGGCCCACTCTGCTGGTGTCTGCTGGCCCTATGCCAAGGAGTCTGGGATC 1003

Db 1021 AGCTTCAAGTTGAAGGCTCAGTGGCTCAAGGACAAACAATTTAGGAGTGCGGTGGTCTGG 1080
Qy 1184 GCCATTGATCGGATGACTTCACTGGCAGCTTCTGCAACAGGGAAGTTTCCCTTAATC 1243
Db 1081 CCCTGACATGGATGACTTCAGTGGTCTTCTGTCACAGGAGCCTTCCCTCTGACA 1140
Qy 1244 TCACCCCTGAAGAGGCCCTCGGCTCGAGAGTGAAGTTGCAAGGCTCCAGCTC 1298
Db 1141 ACTACTTTAAAGAGAGATCTGAATGTACACAGTGAAGTTGCAAGGCTCTTATC 1195

RESULT 15
AR206042
LOCUS AR206042 1656 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 3 from patent US 6372212.
ACCESSION AR206042
VERSION AR206042.1 GI:21504526
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1656)
AUTHORS Gray,P.W.
TITLE Chitinase materials and methods
JOURNAL Patent: US 6372212-A 3 16-APR-2002;
FEATURES
Location/Qualifiers
1..1656
/organism="unknown"
BASE COUNT 365 a 497 c 447 g 347 t
ORIGIN

Query Match 29.7%; Score 483; DB 6; Length 1656;
Best Local Similarity 60.2%; Pred. No. 9.7e-119;
Matches 870; Conservative 0; Mismatches 545; Indels 30; Gaps 3;
Qy 84 GACTGGTGTGACTGCAACATGACAAAGCTTATTCCTCAGAGTCTGTCTATAC 143
Db 7 GCCTGGCGTGTGAGTGCATGTGGTCTGTGGCTGGGAGGTTTCATGTGCTGTC 66
Qy 144 TGAATTTGACGCTGGCTGTGCTTACCAGCTGACATGCTACTTACCAACTGGGCCAGT 203
Db 67 TGATGATCCCATGGGCTCTGTGCAAACTGGTCTGTCTACTTCAACCACTGGGCCAGT 126
Qy 204 ACCGGCCAGCCCTGGGCGCTTCATGCTGACAACTGACCCCTGCTCTGTACCCACC 263
Db 127 ACAGACAGGGGAGGCTCGCTTCTGCCCCAAGGACTTGGACCCAGCCTTTGCACCCACC 186
Qy 264 TGATCTACGCTTGTGGGAGGCGAGCAACAGAGATCACCACTGCAATGGAACGATG 323
Db 187 TCATCTACGCTTGTGGGATGACCAACCAACAGCTGAGCACCAGTGTGGAATGACG 246
Qy 324 TGACTCTTACCAAGCTTTCAATGGGCTGAAATAAGACAGCCAGCTGAAACTCTCC 383
Db 247 AGACTCTTACCAAGGATTTCAATGGGCTGAAGAAGATGAATCCCAAGCTGAAGACCTGT 306
Qy 384 TGGCCATTTGAGGCTGGAATCTGGGAGTCCGCTTCTCACTGCGATGGTTTCTACTCCTG 443
Db 307 TAGCCATCGGAGCTGGAATTTGAGCACTCAGAAGTTTCAAGATATGGTAGCCAGCCCA 366
Qy 444 AGAACCCGACAGCTTTCATCAGCTCAGTATCAAAATCTCTGGCGCAGTATGAGTTGACG 503
Db 367 ACAACCGTACAGCTTGTCAACTCGGCCATCAGGTTTCTGCGCAATATACAGCTTTGACG 426
Qy 504 GGCTGGACTTTGACTGGGAGTACCTTGGCTCTGCTGGGAGCCCTCTCCAGCAAGCATC 563
Db 427 GCCTTGACCTTGTACTGGGAGTACCCAGGAAGCCAGGGAGCCCTCGCGTAGACAGAGC 486
Qy 564 TCTTACTGTCTGGTGCAGGAAATCGTGAAGCTTTTGAAGAGGAGGCCAAGCAGATCA 623
Db 487 GCTTCAACACCTGGTACAGGACTTGGCCAAATGCCTTCCAGAGGAAGCCAGACCTCAG 546
Qy 624 ACAAGCCAGGCTGATGGTGTGCTGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 683

Search completed: July 3, 2003, 07:21:38
Job time : 4281.04 secs

Result No.	Score	Query Match	Length	DB	ID	Description
1	1529.6	94.1	1678	22	AAH42025	Disease treatment
2	1366.4	84.1	1368	22	AAH42013	Disease treatment
3	1266.2	77.9	1369	24	ABW59993	Novel human coding
4	701.8	43.2	1489	22	AAH42023	Disease treatment
5	629.2	38.7	746	22	AAH86935	Human digestive sy
6	628.4	38.7	638	24	ABK11713	RNA encoding novel
7	483	29.7	1656	19	AAV10436	Human chitinase cl
8	483	29.7	1656	20	AAZ21948	MO-13B clone of hu
9	483	29.7	1656	22	AAO33760	Human chitinase cD

DR	WPI; 2001-397791/42.
XX	New proteins, peptides and DNA for treatment of bronchial asthma,
PT	chronic occlusive lung disease and infectious disease -
XX	Claim 5; Page 100; 114pp; Japanese.
PS	The present invention provides the sequence of a protein which can be
CC	used in the treatment and prevention of infectious diseases. Inhibitors
CC	of the protein can be used to treat bronchial asthma and chronic
CC	occlusive pulmonary disease. The present sequence is an oligonucleotide
CC	described in the exemplification of the invention.
XX	
SQ	Sequence 1368 BP; 324 A; 398 C; 349 G; 297 T; 0 other;
	Query Match 84.1%; Score 1366.4; DB 22; Length 1368;
	Best Local Similarity 99.9%; Pred. No. 0;
	Matches 1367; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy	167 TACCAGGTGACATGCTACTTACACAACACTGGGCCAGTCACCGGCCTGGGCGCCTC 226
Db	1 TACCAGGTGACATGCTACTTACCACAACACTGGGCCAGTCACCGGCCTGGGCGCCTC 60
Qy	227 ATGGCTCACAACTCGACCCCTGCCTCTGTACCCACCTGATCTACGGCTTTGCTGGAGG 286
Db	61 ATGGCTGACAACTCGACCCCTGCCTCTGTACCCACCTGATCTACGGCTTTGCTGGAGG 120
Qy	287 CAGAACCAACAGATCACCACTCGAATGGAACGATGTGACTCTCTACCAAGCTTTCAAT 346
Db	121 CAGAACAACAGATCACCACTCGAATGGAATGGAATGATGTGACTCTCTACCAAGCTTTCAAT 180
Qy	347 GCCCTGAAAATAAGAACAGCAGCTGAAACTCTCTGGCCATTGGAGGCTGGAACTTC 406
Db	181.GGCCCTGAAAATAAGAACAGCAGCTGAAACTCTCTGGCCATTGGAGGCTGGAACTTC 240
Qy	407 GGGACTGCCCTTTTCACTGCCCATGGTTTCTACTCTCAGAACCCGACAGCTTTCATCACC 466
Db	241 GGGACTGCCCTTTTCACTGCCCATGGTTTCTACTCTCAGAACCCGACAGCTTTCATCACC 300
Qy	467 TCAGTCATCAAATTCCTCGGCCAGTAGTAGTTTGACGGCTGGACTTTGACTGGGAGTAC 526
Db	301 TCAGTCATCAAATTCCTTGGCCAGTAGTAGTTTGACGGCTGGACTTTGACTGGGAGTAC 360
Qy	527 CCTGGCTCTCGTGGGAGCCCTCTCTCAGGACAAGCATCTCTTCACTGCTCTGGTCAGAGAA 586
Db	361 CCTGGCTCTCGTGGGAGCCCTCTCTCAGGACAAGCATCTCTTCACTGCTCTGGTCAGAGAA 420
Qy	587 ATGGCTCAAGCTTTTGACGAGGAGGCCAACAGAGATCAACAAGCCAGGCTGATGGTCACT 646
Db	421 ATGGCTGAAGCTTTTGACGAGGAGGCCAACAGATCAACAAGCCAGGCTGATGGTCACT 480
Qy	647 GCTGCAGTAGCTGCTGGCATCTCCAATATCCAGTCTGGCTATGAGATCCCCCAACTGTCA 706
Db	481 GCTGCAGTAGCTGCTGGCATCTCCAATATCAAGTCTGGCTATGAGATCCCCCAACTGTCA 540
Qy	707 CAGTACCTGGACTACATCCATGTGATGACCTACGACCTCCATGGCTCTCTGGGAGGGCTAC 766
Db	541 CAGTACCTGGACTACATCCATGTGATGACCTACGACCTCCATGGCTCTCTGGGAGGGCTAC 600
Qy	767 ACTGGAGAGAACAGCCCCCTCTACAAATACCGACTGACACCGGCACGACGCTACCTC 826
Db	601 ACTGGAGAGAACAGCCCCCTCTACAAATACCGACTGACACCGGCACGACGCTACCTC 660
Qy	827 AATGTGTGATTATGTGATGAATCTACTGGAAGGACAAATGGACACCAAGCTGAGAGCTCATC 886
Db	661 AATGTGTGATTATGTGATGAATCTACTGGAAGGACAAATGGACACCAAGCTGAGAGCTCATC 720
Qy	887 GTTGGATTCCCTACCTATGGACACAACTTTCATCTGAGCAACCCCTCCAACTCGGAATT 946
Db	721 GTTGGATTCCCTACCTATGGACACAACTTTCATCTGAGCAACCCCTCCAACTCGGAATT 780
Qy	947 GTTGCCCCCACTCTGTGTGTGTCCTGCTGGGCCCTTATGCCAAGGAGTCTGGAGCTGG 1006

PT encoded polypeptide such as cancer and multiple sclerosis -

XX Claim 1; SEQ ID NO 404; 509pp; English.

XX The present invention provides the protein and coding sequences of 444
CC novel human proteins. These were isolated from expressed sequences tags
CC (ESTs). They can be used to stimulate cell growth, to regulate
CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
CC e.g. in burn treatment, to regulate the immune system e.g. to treat
CC multiple sclerosis, to regulate activin or inhibin e.g. to treat
CC infertility, to regulate haemostasis or thrombolysis e.g. to treat
CC stroke and cancer, to screen for drugs, to treat inflammatory conditions
CC e.g. rheumatoid arthritis, and to treat nervous system disorders e.g.
CC Parkinson's disease. The present sequence is a coding sequence of the
XX invention.

XX Sequence 1369 BP; 327 A; 393 C; 341 G; 308 T; 0 other;

Query Match 77.9%; Score 1266.2; DB 24; Length 1369;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 1274; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 339 CTTTCATGGCCTGAAATAAGACAGCCAGCTGAAACCTCTCTGGCCATTGGAGGCT 398
DB |||||
QY 65 CTTTACTGAATTTGCGAGCTCGGAACAGCCAGCTGAAACCTCTCTGGCCATTGGAGGCT 124
DB |||||
QY 399 GGAACCTTCGGGACTGCCCCCTTTCACCTGCCATGGTTTCTACCTCTGAGAACCGCCAGACTT 458
DB |||||
QY 125 GGAACCTTCGGGACTGCCCCCTTTCACCTGCCATGGTTTCTACCTCTGAGAACCGCCAGACTT 184
QY 459 TCATCAGCTCAGTCATCAATTCCTGGCCAGTATGAGTTTGAGCGGCTGAGCTTTGACT 518
DB |||||
QY 185 TCATCAGCTCAGTCATCAATTCCTGGCCAGTATGAGTTTGAGCGGCTGAGCTTTGACT 244
QY 519 GGGAGTACCCCTGGCTCTGTTGGGAGCCCTCTCAGGACCAAGCATCTCTTCACTCTCTCTGG 578
DB |||||
QY 245 GGGAGTACCCCTGGCTCTGTTGGGAGCCCTCTCAGGACCAAGCATCTCTTCACTCTCTCTGG 304
QY 579 TGCAGGAATTCGTTGAAGCTTTTGGAGGAGGCGCAAGCATCAAGCCCAAGGCTGA 638
DB |||||
QY 305 TGCAGGAATTCGTTGAAGCTTTTGGAGGAGGCGCAAGCATCAAGCCCAAGGCTGA 364
QY 639 TGGTCACTGCTGAGTACCTGCTGGCATCTCCATATCCAGTCTGGCTATGAGATCCCC 698
DB |||||
QY 365 TGGTCACTGCTGAGTACCTGCTGGCATCTCCATATCCAGTCTGGCTATGAGATCCCC 424
QY 699 AACTGTACAGTACCTGAGTACATCCATGCTGAGTACCTGAGTACCTGAGTACCTGAGT 758
DB |||||
QY 425 AACTGTACAGTACCTGAGTACATCCATGCTGAGTACCTGAGTACCTGAGTACCTGAGT 484
QY 759 AGGCTACACTGGAGAGAACAGCCCCCTCTACAAATACCCGACTGACACCGGCGCAACG 818
DB |||||
QY 485 AGGCTACACTGGAGAGAACAGCCCCCTCTACAAATACCCGACTGACACCGGCGCAACG 544
QY 819 CCTACCTCAATGGATTATGCTATGATGATGATGATGATGATGATGATGATGATGATGAT 878
DB |||||
QY 545 CCTACCTCAATGGATTATGCTATGATGATGATGATGATGATGATGATGATGATGATGAT 604
QY 879 AGCTCATGCTGGATTCCCTACCTATGAGACACAACTTCACTCTGAGCAACCCCTCCAAACA 938
DB |||||
QY 605 AGCTCATGCTGGATTCCCTACCTATGAGACACAACTTCACTCTGAGCAACCCCTCCAAACA 664
QY 939 CTGGAATGTGCCCCCACCCTCTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 998
DB |||||
QY 665 CTGGAATGTGCCCCCACCCTCTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 724
QY 999 GGATCTGGGCTTACTAGAGATCTGTACCTTCTTGAAATGAGCCACTCAGGATGGG 1058
DB |||||
QY 725 GGATCTGGGCTTACTAGAGATCTGTACCTTCTTGAAATGAGCCACTCAGGATGGG 784
QY 1059 ATGCCCCCTCAGGAAGTCCCTTATGCCCTATCAGGCAATGTGGTGTGGTGTGGTGTGGTGTGG 1118
DB |||||
QY 785 ATGCCCCCTCAGGAAGTCCCTTATGCCCTATCAGGCAATGTGGTGTGGTGTGGTGTGGTGTGG 844

QY 1119 TCAAGAGCTTCGATATTAAAGCTCAATGGCTTAAGCACAACAAATTTTGAGGCGCATGG 1178
DB |||||
QY 845 TCAAGAGCTTCGATATTAAAGCTCAATGGCTTAAGCACAACAAATTTTGAGGCGCATGG 904
DB |||||
QY 1179 TCTGGGCAATGATCTGGATGACTTCACCTGACCTTTTGTGCAACAGGCGCAAGTTTCCCC 1238
DB |||||
QY 905 TCTGGGCAATGATCTGGATGACTTCACCTGACCTTTTGTGCAACAGGCGCAAGTTTCCCC 964
DB |||||
QY 1239 TAATCTCCACCTCAAGAAAGCCCTCGCCCTGACAGAGTGCAGTTGACGGCTCCAGCTC 1298
DB |||||
QY 965 TAATCTCCACCTCAAGAAAGCCCTCGCCCTGACAGAGTGCAGTTGACGGCTCCAGCTC 1024
DB |||||
QY 1299 AGCCCAATGAGCCCAATAACTGCTCTCCAGTGCAGCGGGAACGGAGGAGTAGCA 1358
DB |||||
QY 1025 AGCCCAATGAGCCCAATAACTGCTCTCCAGTGCAGCGGGAACGGAGGAGTAGCA 1084
DB |||||
QY 1359 GCTCTGAGGAGCAGCTCGGAGGAGTGTGATCTCTGTGTGTCAGAGCAAGCGGCTCTAC 1418
DB |||||
QY 1085 GCTCTGAGGAGCAGCTCGGAGGAGTGTGATCTCTGTGTGTCAGAGCAAGCGGCTCTAC 1144
DB |||||
QY 1419 CCGTGGCAATTAACAGAAATGCCCTTCTGGCACTCGCTGAATGGAGTCACTACAGCAGA 1478
DB |||||
QY 1145 CCGTGGCAATTAACAGAAATGCCCTTCTGGCACTCGCTGAATGGAGTCACTACAGCAGA 1204
DB |||||
QY 1479 ACTGCCAGGCGGCTTGTCTTCGACACCACTGATGCTGATGCTGCAACTGGGCATAAACCT 1538
DB |||||
QY 1205 ACTGCCAGGCGGCTTGTCTTCGACACCACTGATGCTGATGCTGCAACTGGGCATAAACCT 1264
DB |||||
QY 1539 GACCTGTCTATATTCCTTAGAGTTCAGTCTCTTTTGTGTTAGGACATGTTGCCCTTACC 1598
DB |||||
QY 1265 GACCTGTCTATATTCCTTAGAGTTCAGTCTCTTTTGTGTTAGGACATGTTGCCCTTACC 1324
DB |||||
QY 1599 TAAAGTCTCTCAATAAATAACAGCAGTC 1625
DB |||||
QY 1325 TAAAGTCTCTCAATAAATAACAGCAGTC 1351

RESULT 4

AAH42023

ID AAH42023 standard; DNA; 1469 BP.

XX AC AAH42023;

XX DT 24-AUG-2001 (first entry)

XX DE Disease treatment related oligonucleotide SEQ ID NO: 14.

XX KW Disease treatment; infection; chronic occlusive pulmonary disease;

XX KW bronchial asthma; ds.

XX OS Mus sp.

XX PN WO200136633-A1.

XX PD 25-MAY-2001.

XX PF 14-NOV-2000; 2000WO-JP08015.

XX PR 15-NOV-1999; 99JP-0324467.

XX PA (TAKE) TAKEDA CHEM IND LTD.

XX PI Nakanishi A, Morita S;

XX DR WPI; 2001-397791/42.

XX PT New proteins, peptides and DNA for treatment of bronchial asthma,

XX PT chronic occlusive lung disease and infectious disease -

XX PS Disclosure; Page 106; 114pp; Japanese.

XX CC The present invention provides the sequence of a protein which can be

Db 211 GATATTAAAGGCTCAATGGCTTAAGCACAAATAATTGGAGGGCCCGCATGTCTGGGCCATT 270

QY 1190 GATCTGATGACTTCACTGGCACTTTCTGCAACAGGGCAAGTTTCCCTTAATCTCCACC 1249

Db 271 GATCTGATGACTTCACTGGCACTTTCTGCAACAGGGCAAGTTTCCCTTAATCTCCACC 330

QY 1250 CTGAAGAAGGCCCTCGGCCCTGCAGAGTGAAGTTGCACGGCTCCAGCTCAGGCCCAATTGAG 1309

Db 331 CTGAAGAAGGCCCTCGGCCCTGCAGAGTGAAGTTGCACGGCTCCAGCTCAGGCCCAATTGAG 390

QY 1310 CCAATAACTGCTGCTCCAGTGGCAGCGGAGACGGGAGGAGTAGAGCTCTGAGGCG 1369

Db 391 CCAATAACTGCTGCTCCAGTGGCAGCGGAGACGGGAGGAGTAGAGCTCTGAGGCG 450

QY 1370 AGCTGGGAGGAGTGATGCTGCTGTCAGAGCAACGGCTCTACCCCGTGGCAAT 1429

Db 451 AGCTGGGAGGAGTGATGCTGCTGTCAGAGCAACGGCTCTACCCCGTGGCAAT 510

QY 1430 AACAGAAATGCCTTCTGGCACTGCGTGAATGGAGTCACTACCCAGAGAACTGCCAGGCC 1489

Db 511 AACAGAAATGCCTTCTGGCACTGCGTGAATGGAGTCACTACCCAGAGAACTGCCAGGCC 570

QY 1490 GGGCTTGTCTGACACAGCTGTGATGCTGCAACTGGCGGATGAACCTGACTGTCTA 1549

Db 571 GGGCTTGTCTGACACAGCTGTGATGCTGCAACTGGCGGATGAACCTGACTGTCTA 630

QY 1550 TATTCCTAGAGTTCAGTCTCTTTTGTAGGACATGTCGCCCTACCTAAAGTCTCTC 1609

Db 631 TATTCCTAGAGTTCAGTCTCTTTTGTAGGACATGTCGCCCTACCTAAAGTCTCTC 690

QY 1610 AATAAATCAGCAGTC 1625

Db 691 AATAAATCAGCAGTC 706

RESULT 6

ABK11713

ID ABK11713 standard; cDNA; 638 BP.

XX ABK11713;

AC

XX

DT 05-JUN-2002 (first entry)

XX DNA encoding novel human secreted protein #5.

Secreted protein; gene therapy; immunostimulant; human; stroke;
Immune system disorder; immunologic deficiency syndrome; anaemia;
ataxia telangiectasia; HIV; human immunodeficiency virus; heart attack;
Wiskott-Aldrich disorder; thymoblastoma; scarring; erythematosis;
autoimmune disease; Addison's disease; encephalomyelitis; lupus;
multiple sclerosis; autoimmune thyroiditis; diabetes; inflammation;
nephritis; ischaemia; hypergammaglobulinaemia; Crohn's disease;
sarcoidosis; Gaucher's disease; cardiovascular disease; telangiectasia;
persistent truncus arteriosus; arrhythmias; angina pectoris; tumour;
myocardial infarction; hypertension; neovascularisation; angiogenesis;
cancer; ocular angiogenic disease; hyperproliferative disorder;
diabetic retinopathy; uveitis; rheumatoid arthritis; psoriasis;
wound healing; endometriosis; vasculogenesis; atherosclerosis;
nervous system disease; Parkinson's disease; Alzheimer's disease;
infectious disease; gene; ss.

XX Homo sapiens.

XX

XX

PH Key Location/Qualifiers

FT CDS 181..540

FT /tag= a

FT /product= "Novel secreted protein"

FT 181..183

FT sig_peptide /tag= b

FT 184..537

FT mat_peptide /tag= c

FT /label= mature_human_secreted_protein

XX

PN WO200214341-A1.

XX 21-FEB-2002.

XX 13-AUG-2001; 2001WO-US25288.

XX 14-AUG-2000; 2000US-225215P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Barash SC, NI J, Ruben SM, Rosen CA, Shi Y;

XX WPI; 2002-257592/30.

XX P-PSDB; AA077733.

XX Novel human secreted proteins useful for treating immune system e.g.

XX anaemia, cardiovascular e.g. myocardial infarction, neurological

XX disorders e.g. Alzheimer's disease, Parkinson's disease -

XX Claim 1; Page 322; 337pp; English.

XX The invention describes eight novel human secreted polypeptides (I)-(I).

XX and the polynucleotides encoding them are useful for treating, disorders,

XX including disorders of immune system e.g. immunologic deficiency

XX syndrome, ataxia telangiectasia, HIV infection (human immunodeficiency

XX virus), Wiskott-Aldrich disorder, anaemia, thymoblastoma, heart

XX attacks (infarction), strokes, or scarring; autoimmune diseases e.g.

XX Addison's disease, encephalomyelitis, multiple sclerosis, autoimmune

XX thyroiditis, lupus, erythematosis, insulin dependent diabetes mellitus;

XX disorders of inflammation e.g. nephritis, ischaemia-reperfusion injury;

XX hypergammaglobulinaemia, Crohn's disease, sarcoidosis, Gaucher's disease;

XX cardiovascular disease e.g. cardiovascular abnormalities, persistent

XX truncus arteriosus, arrhythmias, angina pectoris, myocardial infarction,

XX hypertension, telangiectasia, ischaemia includes cerebral ischaemia, late

XX mucocutaneous lymph node syndrome; disorders and/or conditions associated

XX with neovascularisation; cancers which involve angiogenesis, benign

XX tumours, ocular angiogenic diseases, hyperproliferative disorders,

XX diabetic retinopathy, uveitis, rheumatoid arthritis, psoriasis, delayed

XX wound healing, endometriosis, vasculogenesis, atherosclerosis; nervous

XX system diseases, degeneration associated with Parkinson's disease, a

XX Alzheimer's disease and infectious diseases. This sequence encodes a

XX novel human secreted protein, described in the invention.

XX

SQ Sequence 638 BP; 153 A; 171 C; 169 G; 145 T; 0 other;

Query Match 38.7%; Score 628.4; DB 24; Length 638;
Best Local Similarity 99.8%; Pred. NO. 5.9e-168;
Matches 629; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 996 CTGGGATCTGGGCTTACTACGAGATCTGTACCTTCTCTGAAAATGGAGCCTCAGGGAT 1055

Db 2 CTGGGATCTGGGCTTACTACGAGATCTGTACCTTCTCTGAAAATGGAGCCTCAGGGAT 61

QY 1056 GGGATCCCCCTCAGGAAGTGCCTTATGCCCTATCAGGGCAATGTGGTGGCTATGACA 1115

Db 62 GGGATCCCCCTCAGGAAGTGCCTTATGCCCTATCAGGGCAATGTGGTGGCTATGACA 121

QY 1116 ACATCAAGAGCTTCGATATTAAGGCTCAATGGCTTAAGCACACAATTTGGAGGGCCCA 1175

Db 122 ACATCAAGAGCTTCGATATTAAGGCTCAATGGCTTAAGCACACAATTTGGAGGGCCCA 181

QY 1176 TGGTCTGGGCAATTTGATCTGGGCTTCTGCTGGGCTTCTGCAACACCAAGTTTC 1235

Db 182 TGGTCTGGGCAATTTGATCTGGGCTTCTGCTGGGCTTCTGCAACACCAAGTTTC 241

QY 1236 CCCTAATCTCCACCTGAAGAGCCCTCGGCTCGCAGTGCAGAGTGCAGGCTCCAG 1295

Db 242 CCCTAATCTCCACCTGAAGAGCCCTCGGCTCGCAGTGCAGAGTGCAGGCTCCAG 301

QY 1296 CTCAGGCCCATTTAGGCCAATAACTGTCTGCCAGTGGCGGAGCGGGAGGAGTA 1355

Db 302 CTCAGGCCCATTTAGGCCAATAACTGTCTGCCAGTGGCGGAGCGGGAGGAGTA 361


```
Db 847 CCTCTCATCAGACACAGAGTGGGGCCCCAGCCACAGGGTCTGGCACTCCAGGCCCT 906
QY 984 ATCCAGGAGTCTGGGATCTGGCTTACTACAGATCTGTACTTCTCTGAAATGGAG 1043
Db 907 TCACCAAGGAGGAGTCTGGCTTACTATGAAGTCTCTCTCTCTCTCTCTCTCTCT 960
QY 1044 CCACCTAGGAGTGGGATGCCCTCAGGAAGTGGCTTATGCTATCAGGGCAATGTGTGG 1103
Db 961 CCACCAACAGAGAAATCCAGGATCAGAAGGTGCCCTACATCTTCCGGGACACCAAGTGG 1020
QY 1104 TTGGCTATGACACATCAAGAGCTTCGATATTAGGCTCAATGGCTTAAGCAACAAT 1163
Db 1021 TGGGCTTTGATGATGGAGAGCTTCAAAACCAAGGTACGCTATCTGAAGCAGAAGGAC 1080
QY 1164 TTGGAGGCGCATGGTCTGGGCCATTGATCTGGATGACTTCACTGGCACTTTCTGCAACC 1223
Db 1081 TGGGCGGGCGCATGGTCTGGGCACTGGACTTAGATGACTTGGCGGCTTCTCTGCAACC 1140
QY 1224 AGGCAAGTTTCCCTAATCTCCACCTGAAGAGCCCTCGGCCCTGCAGAGTCAAGTT 1283
Db 1141 AGGCGGATACCCCTCATCCAGACGCTACGGCAGGAAGTGAAGTCTTCCATCTTGCCTT 1200
QY 1284 GCAGGCTCAGCTACGCCATTTAGCCAAATAGTGTCTCCAGTGGCAGCGGGAAG 1343
Db 1201 CAGGCAACCCAG---AGCTTGAAGTTCCAAACCAAGGTACAGCCCTCTGAACCTGAGCA-- 1255
QY 1344 GGAGCGGAGTAGCAGCTCTGGAGGCGCTCGGGAGGCGAGTGATCTGTCTGCTCAGAG 1403
Db 1256 -----TGGCCCCAGCCCTGGACAAGACAGCTTCTGCCAGGGGCAAG 1296
QY 1404 CCAACGGCTCTACCCCGTGGCAAAATACAGAAATGCTTCTGGCACTGGGTGAATGGAG 1463
Db 1297 CTGATGGCTCTATCCCAATCTCGGGAAGGTCCAGCTTCTACAGCTGTGCAGCGGGC 1356
QY 1464 TCAGCTACAGAGAGACTGCCAGCGGGGTTCTCTTCGACACCAAGTGTGATGTGCA 1523
Db 1357 GCGTGTTCAGCAAGCTGCGCCAGACAGGCTGTGTTCAGCAACTCTGCAATGTGCA 1416
QY 1524 ACTGG 1528
Db 1417 CCTGG 1421
```

RESULT 8

```
AAZ21848
ID AAZ21848 standard; DNA; 1656 BP.
XX
AC AAZ21848;
AT
DT 10-DEC-1999 (first entry)
XX
DE MO-13B clone of human Chitinase, with noncoding 5'/3' regions.
XX
KW chitin; fungal infection; immunocompromised; AIDS; chemotherapy;
KW organ transplant; parasite; chitin-binding; allele; vector;
KW truncated protein; ds.
XX
OS Homo sapiens.
```

```
Key Location/Qualifiers
FT CDS 27..1427
FT /tag= a
FT /product= Human_Chitinase
FT sig_peptide 27..89
FT /tag= b
FT /note= "Signal peptide"
FT mat_peptide 90..1427
FT /tag= c
FT /note= "Mature peptide"
```

W09946390-A1.

16-SEP-1999.

```
XX
PF 12-MAR-1999; 99WO-US05343.
XX
PR 12-MAR-1998; 98US-0039198.
XX
PA (ICOS-) ICOS CORP.
XX
PI Gray PW, Tjoelker LW;
XX
XX WPI; 1999-551417/46.
DR P-PSDB; AAY42426.
XX
PT Novel chitin-binding fragments of human chitinase used to treat fungal
XX infections in animals
XX
PS Example 1; Page 59-62; 83pp; English.
XX
CC This is the nucleotide sequence of an allelic form of the human
CC chitinase enzyme, which is capable of degrading Chitin (a linear
CC homo polymer of beta-1,4-linked N-acetylglucosamine residues).
CC Chitinase fragments can be used to screen for proteins or other
CC molecules that specifically bind to the chitin-binding domain of human
CC chitinase or that modulate its activity. These compounds are useful for
CC immunization, as well as for purifying chitinase. Polynucleotide fragments of
CC detection and quantification of chitinase. Polynucleotide fragments of
CC the invention are useful as a source of probes and primers, and to
CC express the proteins recombinantly. The chitinase fragments, when
CC conjugated to antifungal compounds, are used to treat animals,
CC especially humans, infected with chitin-containing parasites such as
CC fungi. Fungal infection treated include candidiasis, aspergillosis,
CC coccidioidomycosis, blastomycosis, paracoccidioidomycosis,
CC mucormycosis, histoplasmosis, cryptococcosis, chromoblastomycosis,
CC sporotrichosis, and dermatophytoses.
CC Chitin can be degraded by the enzyme chitinase. Use of whole chitinase
CC protein for treating infections, especially fungal infections, is
CC problematic. In view of the increasing incidents of life-threatening
CC fungal infection in e.g. immunocompromised individuals, there exists a
CC need for identifying new compounds for treating fungal infection. The
CC chitin-binding fragments of the present invention provide this need.
XX
SQ Sequence 1656 BP; 365 A; 497 C; 447 G; 347 T; 0 other;
```

```
Query Match 29.7%; Score 483; DB 20; Length 1656;
Best Local Similarity 60.2%; Pred. No. 1.8e-126;
Matches 870; Conservative 0; Mismatches 545; Indels 30; Gaps 3;

QY 84 GACTGTGTGCTGACTGCAACCATGACAAAGCTTATCTCTCACAGGCTTTGTCTTATAC 143
Db 7 GCCTGCCGTGAGCTGCATCATGCTGCGGTCTGTGGCTGGCAGGTTTCATGGTCTCTG 66

QY 144 TGAATTTGCAGCTCGGCTCTGCCTACCAGTGCATGCTACTTCCACCACTGGGCCAGT 203
Db 67 TGATGATCCCATGGGCTCTGCTGCAAAACTGTCTGCTACTTCCACCACTGGGCCAGT 126

QY 204 ACCGGCCAGGCTCTGGGCGCTTTCATGCTGCACAAATCGACCCCTGCTGTGACCCACC 263
Db 127 ACAGACAGGGGAGGCTCGCTTCTGCCCAAGACTTGGACCCAGGCTTGGCACCACC 186

QY 264 TGATCTACGCTTTTGTGGGAGCGAGAACCAACAGATACACCATCGAATGAAGCATG 323
Db 187 TCATCTACGCTTTCGCTGGCATGACCAACCCAGCTGAGCAGCTAGTGAATGACG 246

QY 324 TGACTCTTACCAAGCTTTCATGGCTGGAATAATGAACACACCCAGCTGAAACCTCC 383
Db 247 AGACTCTTACCAAGGAGTTCAATGGCTTGAAGAAGATGAATCCCAAGCTGAAGACCTGT 306

QY 384 TGGCCATTTGGAGGCTTGAAGTCTCGGAGTGCCTTCTGCTGCTGCTGCTTCTTCTCTCT 443
Db 307 TAGCCATTTGGAGGCTTGAATTTTCCAGCACTCAGAAATTCACAGATATGTTAGCCAGG 366

QY 444 AGAACCGCCAGACTTTTCATCACCCTCAGTCAATCAATTCCTGCGCCAGTATGATGACG 503
Db 367 ACAACCGTCAAGCTTTGTCACTCGGCCATCAGGTTTCTGCGCAATAACAGCTTTGACG 426
```

RESULT	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	524	525	526	527	528	529	
--------	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	--

84 GACTGGTGCTGACTGCACCCATGACAAAGCTTATTCTCCTCACAGGTCCTTGTCTTATAC 143

84 GACTGGTGCTGACTGCACCCATGACAAAGCTTATTCTCCTCACAGGTCCTTGTCTTATAC 143

Db 7 GCCTGCCGTGAGCTGCATCATGTGCGGTCTGTGGCCTGGGCGAGGTTTCATGGTCTGCG 66
QY 144 TGAATTTGAGCTCGGCTCTGCTACCAAGCTGACATGCTACTTACCACACTGGGCCCCAGT 203
Db 67 TGTATGATCCCATGGGCTCTGTGCAAACTGGTCTGTACTTTCACCAACTGGGCCCCAGT 126
QY 204 ACCGCCAGGCTGGGGGCTTCATGCTCTGACAACTGACCCCTGCTCTGTACTACCCACC 263
Db 127 ACAGACGGGGAGGCTGCTCTGCTGCCAGGACTTGGACCCAGCCCTTGCACCCACC 186
QY 264 TGTATAGCCCTTCTGCGGAGGAGAGACAGAGATCACCACCATCGAATGGAAGCATG 323
Db 187 TCATCTAGCCCTTCTGCTGGCATGACCAACCAAGCTGAGCAGCCTGAGTGAATGAGC 246
QY 324 TGACTCTCTACCAAGCTTCAATGCTGAAATTAAGAACAGCAGCTGAAACCTCTCC 383
Db 247 AGACTCTCTACCAAGGTTCAATGGCTGGAAGATGAATCCCAAGCTGGAAGCCCTGT 306
QY 384 TGGCCATTGGAGGCTTGGAGCTGCGCCCTTTCACCTGCCATGGTTTCTACTCCTG 443
Db 307 TAGCCATCGGAGGCTGGAATTTACGACTCAGCAATTCACAGATATGTTAGCCAGGCA 366
QY 444 AGAACGCCAGACTTTCATACCTCAGTCATCAATTCCTGGGCGAGTATGATTTGAGC 503
Db 367 ACAACCGTCAGACCTTTGTCACTCGGCCATCAGGTTTCTGGCAAAATACAGCTTTGAGC 426
QY 504 GGCCTGACTTTCAGCTGGGAGTACCCTGGCTCTGCTGGGAGCCCTCTCAGGACAGCATC 563
Db 427 GCCTTGACCTTACCTGGGAGTACCCAGGAAGCAGGGGAGCCCTGCCGTAGACAGGAGC 486
QY 564 TCTTCACCTGCTCCTGGTGGCAAAATCGGTGAAGCTTTTGGAGGAGGCCCAAGAGATCA 623
Db 487 GCTTCACACCTGCTGACAGGACTTGGCCAAATGCTTCAGCAGGAAGCCAGACCTCAG 546
QY 624 ACAAGCCAGGCTGATGCTGCTGCTGAGTGTGCTGGCATCTCCAAATATCCAGTGTG 683
Db 547 GGAAGGAGCGCTTCTTCTGAGTGCAGCGGTTCCAGCTGGGCGAGACCTATGTGGATGTG 606
QY 684 GCTATGAGATCCCAACTGTACACAGTACCTGAGTACATCCATGCTGACCTACGACCT 743
Db 607 GATAGAGGTGACAAATTCGCCAGAACCTCGATTTGTCAACCTTATGGCTACGACT 666
QY 744 TCCATGGCTCTGGAGGCTTACACTGGAGAGAACGCCCTCTACAAATACCCGACTG 803
Db 667 TCCATGGCTCTGGAGAGGCTACCGGACATACAGCCCTCTACAGAGAGCAAGAG 726
QY 804 ACACGGCAGCAAGCCTACCTCAATGTGGATATGCTATGAATCTGGAAGAGCAATG 863
Db 727 AGAGTGGTGCAGCAGCCAGCCCTCAACGTGGATGCTGTGCAACAGTGGCTGCAGAGG 786
QY 864 GAGCACCAGCTGAGAGGCTCATCTGTTGATTCCTTACCTATGACACACTTCATCTGA 923
Db 787 GGACCCCTGCCAGAGCTGATCTTGGATCCCTACCTTACGAGCAGCTCTTACACTGG 846
QY 924 GCAACCCCTCCAAACTTGAATTTGGTGGCCCACTCTGCTGCTGCTCTGCTGGGCCCT 983
Db 847 CCTCTCATCAGACACAGAGTGGGGCCCCAGCCACAGGGTCTGGCACTCCAGGCCCT 906
QY 984 ATGCCAAGAGTCTGGGCTTACTAGAGATCTGTACCTTCTCTGAAAATGAGG 1043
Db 907 TCACCAAGGAGGAGGAGTCTGGCTTACTATGAATCTGCTCTCTCTCTCTCTCTCTCT 960
QY 1044 CCACCTCAGGATGGGATGCCCTCAGGAAGTGCCTTATGCCCTATCAGGCAATGTGGG 1103
Db 961 CCACCAACAGAGAATCCAGGATCAGAGGTGCCCTTACATCTTCCGGGACACCACTGGG 1020
QY 1104 TTGGCTATGACAACTCAAGAGCTTCGATATTAAGGCTCAATGGCTTAAGCAACAAT 1163
Db 1021 TGGGCTTGTATGATGTGGAGACTTCAAAACCAAGTCACTATCTGAAGCAGAGAGGAC 1080
QY 1164 TTGGAGGCGCCATGCTGGGCCATGATCTGGATGACTTCACTGGCACTTTCTGCAACC 1223
Db 1081 TGGGCGGGGCCATGCTTGGGCACTGGACTTAGATGACTTTGCCGCTTCTCTCTCTCTCTCT 1140

QY 1224 AGGCAAGCTTTCCCTAATCTCCACCTGAAGAAGSCCTCGGCTCGAGAGTGCAGATT 1283
Db 1141 AGGGCCGATACCCCTCATCCAGAGCTGACGCGAGAACTGAGTCTTCCATACTTGGCTT 1200
QY 1284 GCACGGCTCCAGCTCAGCCCATTTAGCCAAATAACTGCTGCTCCAGTGGCAGCGGAAG 1343
Db 1201 CAGGCACCCAG---AGCTTGAAGTTCCAAAACCAAGGTGAGCCCTCTGAACCTGAGCA-- 1255
QY 1344 GGAGCGGAGTAGCAGCTCTGGAGCAGCTCGGGAGGAGTGGATTTCTGCTGTCAGAG 1403
Db 1256 -----TGGCCCGAGCCCTGGCAAGACACAGTCTTCTGCCAGGGCAAG 1296
QY 1404 CCAACGGCTCTACCCGTGGCAAAATAACAGAAATGCCTTTCTGGCACTGCGTGAATGGAG 1463
Db 1297 CTGATGGCTCTATCCCAATCTCGGAGCGTCCAGCTTCTACAGCTGTGCGAGGGGC 1356
QY 1464 TCACGTACCAAGCAACTGCCAGGCGGGCTTGTCTTCACACCAAGCTGTGATGCTGCA 1523
Db 1357 GSCCTTCCAGCAAGCTGCCCGACAGGCTGTGTTCAGCAACTCTCTGCAAAATGCTGCA 1416
QY 1524 ACTGG 1528
Db 1417 CCTGG 1421

RESULT 10

ABL57381

ID ABL57381 standard; cDNA; 1656 BP.

XX ABL57381;

AC ABL57381;

DT 12-AUG-2002 (first entry)

XX Human chitinase cDNA clone MO-13B.

DE Chitinase; enzyme; human; fungicide; antifungal; infection;

KW candidiasis; aspergillosis; coccidioidomycosis; blastomycosis;

KW paracoccidioidomycosis; histoplasmosis; cryptococcosis;

KW chromoblastomycosis; sporotrichosis; mucormycosis; dermatophytosis;

KW Pneumocystis; gene; ss.

XX Homo sapiens.

XX Key

FH Location/Qualifiers

CDS 27..1427

FT /*tag= a

FT /product= "Chitinase"

FT 27..89

FT /*tag= b

FT mat_peptide 90..1424

FT /*tag= c

XX US6372212-B1.

XX 16-APR-2002.

XX 16-JUN-1997; 97US-0877599.

XX 14-JUN-1996; 96US-0663618.

XX (ICOS-) ICOS CORP.

XX Gray PW;

XX WPI; 2002-442449/47.

XX P-PSDB; ABB76292.

XX Co-administering chitinase to improve the effectiveness of fungicidal

XX drugs e.g. amphotericin B or itraconazole, useful for treating fungal

XX infections e.g. candidiasis, coccidioidomycosis and blastomycosis -

XX Example 1; Column 31-34; 26pp; English.

Qy		584	GCTATGAGATCCCCCAACTGTCACAGTACTGGCTGCATACCTCATGTCTACGACTCAGACC	743
Db		607	GATACAGGTGGCAAATAATGCCCAAGACTGGATTTTGTCAACCTTATGGCTACGACT	666
Qy		744	TCCATGGCTCCTGGAGGGCTACACTGGAGACAAGCCCCCTCTCAAAATACCGCACTG	803
Db		667	TCCATGGCTCTTGGGAGAAGGTCACGGGACATTAACAGCCCCCTCTCACAGAGCGAAGAAG	726
Qy		804	ACACCGSCAGCAACGCTTACTCAATGTGGATTATGTCAATGAATCTACTGGAAGGACAAATG	863
Db		727	AGAGTGTGTGCAGCAGCAGCCCTCAACGTGGATGCTGTGTCAACAGTAGTGCTCGACAAGG	786
Qy		864	GAGCACCACTGAGAACGCTCATCGTTGGATTCCCTACCTATGGACACAACCTTCACTCTGA	923
Db		787	GGACCCCTGCCAGCAAGCTGATCCTTGGCATGCTTACTACGGACGCTCCTCTCACACTGG	846
Qy		924	GCAACCCCTCCAACACTGGAATTGGTGGCCCCCACTCTTGGTGGTGGCTCTGGCTGGGCCCT	983
Db		847	CCTCCTCATCAGACACCCAGAGTGGGGSCCCCAGCCACAGGGTCTGGCACTCCAGGCCCT	906
Qy		984	ATGCCAGGAGCTGGGACTGGGCTTACTACGAGATCTGTACTCTCTGAAAATGGAG	1043
Db		907	TCACCAAGGAAGAGGGATGCTGGCTTACTATGAAGTCTGCTCCT-----GGAAGGGGG	960
Qy		1044	CCACTCAGGGATGGGATGCCCTTCAGGAAGTGCTTATGCTTATCAGGGCAATGTGTGGG	1103

Query Match	29.7%	Score 483;	DB 24;	Length 1656;
Best Local Similarity	60.2%	Pred. NO. 1.8e-126;		
Matches 970.	Casconative	0.0	Microbial	546

Qy	1164	TTGGAGCGCCATGGTCTGGGCCATGTGATCTGATGACTTCACCTGGCACTTCTCTGCAACC	1222
Db	1081	TGGGCGGGGCATGGTCTGGGCACCTTAGATGACTTTGCGGCGTCTCTCTGCAACC	1140
Qy	1224	AGGCGAAGTTTCCCTTAATCTCCACCCTGAAGAGGCCCTCGGCGCTGCAGGTGCAAGTT	1283
Db	1141	AGGCGCGATACCCCTCATCCAGCGCTACGGCAGGAAGTGTCTTCACATCTTGCCCTT	1200
Qy	1284	GCACGGCTCCAGCTCAGCCCATTTAGGCCAATAACTGCTGCCAGTGGCAGCGGAAACG	1343
Db	1201	CAGGCACCCAG--AGCTTGAAGTTCCAAAACCAAGTCAGCCCTCTGAACCTGAGCA--	1255
Qy	1344	GGAGCGGAGTAGCAGCTCTGGAGGCACTTCGGGAGGCACTGGAATCTGTGCTGTGACAG	1403
Db	1256	-----TGGCCCCAGCCCTGGACAAGACAGCTTCTGCGCAGGGCAAG	1296
Qy	1404	CCAACGGCCTTACCCCGTGGCAATAACAGAAATGCCTTCTGGCACTGCGTGAATGGAG	1463
Db	1297	CTGATGGGCTCTATCCCAATCTCTGGGAGCGGTCCAGCTTCTACAGCTGTGCAGCGGGC	1356
Qy	1464	TCACGTACCAAGCAGAACTGCGACGCGGGCTTCTTCGCACACCACTGTGATTTGCTGCA	1523
Db	1357	GGCTGTTCCAGCAAGCTGCCCCACAGGCGTGTGTTGACCAACTCCTGCAATGCTGCA	1416
Qy	1524	ACTGG 1528	
Db	1417	CCTGG 1421	
RESULT 11			
AAT89181			
ID	AAT89181 standard; cDNA; 1768 BP.		
XX	AAT89181;		
XX	27-APR-1998 (first entry)		
XX	Human chitotriosidase variant cDNA.		
XX	Chitotriosidase; tissue remodelling disorder; diagnosis; therapy;		

RESULTS II
AAT89181
ID AAT89181 standard; cDNA; 1768 BP.

AA
DT 27-APR-1998 (first entry)

Human chitotriosidase variant cDNA.

AA Chitotriosidase; tissue remodelling disorder; diagnosis; therapy;
KW
KW
KW rheumatoid arthritis; atherosclerosis; human; ss.

XX OS Homo sapiens.
 XX FH Key Location/Qualifiers
 XX FT CDS 124..1524
 XX FT /*tag= a
 XX PN W09736917-1.
 XX PN 09-OCT-1997.
 XX PD
 XX PF 21-MAR-1997; 97WO-US05072.
 XX PR 29-MAR-1996; 96US-0014295.
 XX PA (HUMA-) HUMAN GENOME SCI INC.
 XX PA (SMIK) SMITHKLINE BEECHAM CORP.
 XX PI Chopra A, Gentz R, Kirkpatrick RB, Ni J, Thotakura NR;
 XX PI WPI; 1997-503041/46.
 XX DR P-PSDB; AAW31498.
 XX PT New isolated human chitotriosidase gene - used to develop products
 XX PT for the diagnosis and treatment of tissue remodeling disorders, e.g.
 XX PT rheumatoid arthritis
 XX PS Claim 1; Page 21-22; 34pp; English.
 XX CC This human cDNA sequence encodes chitotriosidase (see AAW31498). It
 CC is a preferred variant of the chitotriosidase cDNA sequence given
 CC in AAW89180. Also claimed are: (1) a method of diagnosing a tissue
 CC remodelling disorder related to expression of a mutated
 CC chitotriosidase protein in a host comprising carrying out nucleic
 CC acid amplification; and (2) a method of detecting altered
 CC expression of a chitotriosidase protein in a host comprising
 CC contacting a bodily sample with an antibody. Chitotriosidase can
 CC degrade extracellular matrix substrates with a similar carbohydrate
 CC structure to chitin. It can be used to develop products which can
 CC be used in the diagnosis and treatment of tissue remodeling
 CC disorders such as rheumatoid arthritis or atherosclerosis.
 CC Chitotriosidase nucleic acids can be used in the recombinant
 CC production of the enzyme, and as probes or primers in diagnostic
 CC assays.
 XX SQ Sequence 1768 BP; 401 A; 529 C; 474 G; 363 T; 1 other;
 Query Match 29.7%; Score 482; DB 18; Length 1768;
 Best Local Similarity 60.2%; Pred. No. 3.6e-126;
 Matches 866; Conservative 1; Mismatches 541; Indels 30; Gaps 3;
 QY 91 GCTGACTGCAACCATGACAAAGCTTATTCCTCAGAGTCTGTCTTATCTGATTTT 150
 DB 111 GGTGAGCTGCATCATGGTGGCGTCTGTGGCCTGGGAGGTTTCATGCTCTGATGAT 170
 QY 151 GCAGCTCGGCTCTGCCTACAGCTGACATGCTACTTACCACTGGGCCAGTACCGGCC 210
 DB 171 CCCATGGGCTCTGCTGAAACTGGTCTGCTACTTACCACTGGGCCAGTACAGACA 230
 QY 211 AGGCTTGGGCGCTTCATGCCTGACAACTGCAACATCGACCCCTGCTGTGACCCATCTA 270
 DB 231 GGGGAGGCTCGCTTCTGCCAAGGACTTGGACCCAGCCTTTTGACCCACCTCATCTA 290
 QY 271 CGCCTTGTGGGAGCAACAACAGATACCACTACCACTCGAATGGAAGATGTACTCT 330
 DB 291 CGCCTTGTGGATGACCAACACCACTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 350
 QY 331 CTACCAAGCTTTCAATGGCTGAAATAAGAACACCCAGCTGAAACTCTCTGCGCAT 390
 DB 351 CTACCAAGAGTTCAATGGCTTGAAGAAGATGAATCCAAAGCTGAAGACCTGTGACCAT 410
 QY 391 TGGAGGCTGGAATCTGGGAGTGCCTTTTCACTGCAATGGTTTCTACTCTCTGAGAACCG 450
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

DB 411 CGGAGGCTGGAATTTTCRGCACTCAGAAAGTTTACAGATATGGTAGCCACGCCAACACCG 470
 QY 451 CCAGAGCTTTTCATCACCTCAGTCAATAATTTCTCGGCCAGTATGAGTTTTCAGCGGCTGGA 510
 DB 471 TCAGACCTTTGTCAACTCGCCATCAGGTTTCTGCGCAATAATACAGCTTTTACAGCGCTTGA 530
 QY 511 CTTTGACTGGGAGTACCTTGGCTCTCGTGGGAGCCCTCCTCAGGACAAGCATCTCTTTCAC 570
 DB 531 CCTTGACTGGAGTACCCAGGAAGCCAGGGAGCCCTGCGGTAGACAAGAGGCGCTTCAC 590
 QY 571 TGTCTGTGTGAGGAATAATCGGTGAAGCTTTTTCAGGAGGAGGCAACAGCATCAACAAGCC 630
 DB 591 AACCTTGTGTACAGGACTTGGCCAAATGCTTCCAGCAGGAAGCCAGACCTTCAGGGAAGA 650
 QY 631 CAGGCTGATGCTCAGTGTGAGTGTGATCTCCATATCTCCATATCAGTCTGCTGGCTATGA 690
 DB 651 AGCCTTCTCTGTGAGTGCAGCGGTTCCAGCTGGGAGACCTATGTGGATGCTGGATACGA 710
 QY 691 GATCCCCCAACTGTTCACAGTACCTGAGTACATCTCATGTATGACCTAGACCTCCATGG 750
 DB 711 GGTGGACAATAATCGCCAGAACCTGGATTTTGTCAACCTTATGGCCTACGACTTCATGG 770
 QY 751 CTCCTGGGAGGCTACACTGGAGAGAACACCCCTCTACAAATACCCGACTGACACCGG 810
 DB 771 CTCCTGGAGAGGTCACGGGACATAACAGCCCTCTACAGAGGCAAGAGAGAGTGG 830
 QY 811 CAGCAACGCTACCTCAATGTGATTTATGTCACTACTGGAAGACAATGGACAC 870
 DB 831 TGCAGCAGCAGCCTCAACGTGGATGCTGTGTCAACAGTGGCTCAGAAGGGGACCCC 890
 QY 871 AGCTGAGAAGCTCATCGTTGGATTCCTTACCTATGGACACAACTTTCATCTAGCAACCC 930
 DB 891 TGCCAGCAGCTGATCTTGGCATGCTTACCTACGAGAGCTCTTTCACACTGCGCTCCTC 950
 QY 931 CTCCAACACTGGAATTTGGTGGCCCTACCTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 990
 DB 951 ATCAGACAGCAGAGTGGGGGCCACAGGCTGTGCACTTCCAGGCGCTTCCACCA 1010
 QY 991 GGAGTCTGGAGTGGGCTTACTACGAGATCTGTACTTCTCTGAAAATGGAGCCACTCA 1050
 DB 1011 GGAAGAGGAGTGTGCTGCTTACTATTAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1064
 QY 1051 GGGATGGGATGCCCCCTCAGGAAGTGCCTTATGCTATCAGGCAATGTGTGGTGTGGCTA 1110
 DB 1065 ACAGAAATCCAGGATCAGAAGTGGCTTACATCTTCCGGGACAACAGTGGTGGGCTT 1124
 QY 1111 TGACAACATCAAGAGCTTCGATATTAAGGCTCAATGGCTTAAGCAACAATAATTTGGAG 1170
 DB 1125 TGATGATGTGGAGAGCTTCAAAACCAAGGTTCAGCTATCTGAAGCAGAAGAGGACTGGCGG 1184
 QY 1171 CGCCATGCTGGGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1230
 DB 1185 GGCATGGTCTGGGCTGAGCTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1244
 QY 1231 GTTTCCTCTTAATCTCCACCTTGAAGAAGCCCTCGGCTTGCAGAGTGCAGAGTGCAGGCT 1290
 DB 1245 ATACCCCTTCATCCAGCCTAGCGGAGGAAGTGTCTTCCATATCTGCTTTCAGGCTAC 1304
 QY 1291 TCAGCTAGCCCATTTAGCCCAATAACTGTGTCTCCAGTGGCAGGAGGCAACGGGAGCGG 1350
 DB 1305 CCAG---AGCTTGAAGTTCCAAACCAAGGTACAGCTCTGAACTCTGAGCA----- 1352
 QY 1351 GAGTACAGCTCTGGAGGAGCAGTGGGAGGAGTGTGATCTGTGCTGTCTGAGAGCCACCG 1410
 DB 1353 -----TGGCCCCAGCCCTGGACAAGACAGTGTCTGCGAGGGCAAGAGCTGATGG 1400
 QY 1411 CTTCTACCCCTGGCAATAAACAAGTGTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1470
 DB 1401 GCTCTATCCCAATCTTCGGGAAGCTTCCAGCTTCTACAGTGTGTGACAGCGGGCGGTGTT 1460
 QY 1471 CCAGCAGAACTGCCAGCGCGGCTTGTCTTTCGACACCACTGATGTGATTTGCTGCAACTGG 1528
 DB 1461 CCAGCAAGCTGCCCGCAGGCGCTGTGTGTTTACGCAACTCTCTGCAAAATGCTGCACCTGG 1518

RESULT 12

REGOLI 12
AAT50833

AAT50833
ID AAT50833 standard: cDNA: 1643 bp

AA130833

XX
AC
BATT50833.

AA150833; AC YY

. 24-MAR-1997 (first entry)
Human chitinase cDNA clone chi.50.

XX Chitinase; chitotriosidase; chitin; infectious disease;
KW gene therapy; vaccine; lysosomal lipodosis; Gaucher's disease;
KW leishmaniasis; sarcoidosis; X-linked adrenoleukodystrophy;
KW multiple sclerosis; drug delivery; cosmetics; food; ss.
XX

XX
XX
XXXXX-FFFFF

OS Homo sapiens.

[illegible]

Key	Location
FH	

13...1413|

```
ET  
FT  
CDS  
IS..I413|  
/*tag= a
```

FT
sia pentide

IT, and the

FT	mat	pentide	/lag=	D
76	1410			

```
FI mac_peptide
FT /0::1410|
/*tag=0
```

	FI	YY
/ *lag=		

YY
DN
W09610940-8

PN WO9640940-A2.
yy

XX
10-REQ-1006

PD	19-DEC-1996.	
XX		
XX	06-JUN-1996;	96WO-NL00225.
XX		
XX	07-JUN-1995;	95US-0486839.
XX		
XX	(UNAM) UNIV AMSTERDAM.	
PA		
XX	Aerts JMFG;	
XX		
XX	WPI; 1997-118698/11.	
DR	P-PSDB; AAW08584.	
DR		
XX		
XX		
XX		
PT		New human chitinase - used to treat or prevent infection by
PT		chitin-contg. pathogens, in diagnosis and as additives to cosmetics,
PT		foods, implants etc.

XX
:000 COUNTDOWN / ESCORT T-7

Claim 2; Page 39-40; 58pp; English.

A cDNA clone (AA7508333), designated chi.50, codes for a 50 kDa human chitinase (AAW08584) that is stable to many proteases, active at pH 3-8 and up to 50 deg, and stable in the circulation. Clones chi.50 and chi.39 (see also AA750834) were isolated from a human macrophage cDNA library using as probe a partial clone obtd. using primers (see also AA750835-36) based on a chitotriosidase purified from a type 1 Gaucher disease patient. The 2 cDNA clones are the result of alternative splicing of RNA. Chitinase nucleic acid can be used for large-scale prodn. of recombinant human chitinases, or can be incorporated into a gene therapy vector to treat or prevent infection by chitin-contg. pathogens.

[illegible]

Sequence 1643 BP: 364 A: 490 C: 442 G: 347 T: 0 other:

[illegible]

Query Match

Query Match 25.0%, Score 461.4, DB 16, Length 1043;
Best Local Similarity 60.38: Pred. No. 5 1e-126:

Best local similarity	Fied. NO: 3.1E-120;	
Matches 866: Conservative	0: Mismatches 541:	Indels 30: Cans 3

Matches 500, conservative 541; Indels 30; Gaps 3

09 02 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041

QY	212	GGCTGGGGCGCTTCATGCTGTACACATCGACCCCTGCTCTGTACCCACCTGATCTAC	271
Db	121	GGGAGAGCTGGCTTCCTGCCCCAAGGAGCTTGGAGCCCGACGCTTTGGACCCACCTCATCTAC	180
QY	272	GCCTTTGCTGGAGGCAACAACGAGATCACCACCATCGAATGGAAAGATGTGACTCTC	331
Db	181	GCCTTCCTGGCATGACCAACACGCTGAGCACCACTGAGTGGAAATGACGAGACTCTC	240
QY	332	TACCAAGCTTTCAATGGCCCTGAANAATAAGAACAGCCAGCTGTAATACTCTCTGGCCATT	391
Db	241	TACCAAGAGTTCAATGGCCCTGAAGAAGATGAATCCCAAGCTGAAGACCCCTGTTAGCCATC	300
QY	392	GGAGGCTTGGAACTTCGGAGCTGCCCTTTTCACCTGCCATGGTTTCTACTCCTGAGAACCGC	451
Db	301	GGAGGCTTGGAACTTCGGCATCTCAGAAGTTCACAGATATGGTAGCCAGGCCAACAAACCGT	360
QY	452	CAGACTTTTCATACCTCAGTCATCAAAATTCCTGGCCAGTATGAGTTTTCAGCGGCTGGAC	511
Db	361	CAGACCTTTGTCAACTCGGCCATCAGGTTTTCGGCAAAATACAGCTTTGACGGCCCTTGAC	420
QY	512	TTTGACTGGAGATACCTTGGCTCTGCTGGAGAGCCCTCCTCAGACAAAGCATCTCTTCACT	571
Db	421	CTTGACTTGGAGTAGCCACGAAGCCAGGGGAGGCCCTGCCGTAGACAAGGAGCGCTTCACA	480
QY	572	GTCCTGGTGCAGGAATTCGTGAAGCTTTTTCGACGAGGAGCCCAAGCAGATCAACAACGCC	631
Db	481	ACCTGTGTACAGNCTTGGCCATGCTCTCCAGCAGGAAGCCACAGACCTCAGGGAAGGAA	540
QY	632	AGGCTGATGTCACCTGCTGTCAGTAGCTGGTGGCATCTCCAAATATCCAGTCTGTCATGAG	691
Db	541	CGCCTTCTTGAGTGCAGCGGTTCCAGCTGGGCAGACCTATCTGGATGCTGGATACGAG	600
QY	692	ATCCCCAACTGTCACAGTACCTGGACTACATCCATGTCATGACCTACGACCTCCATGGC	751
Db	601	GTGGACAAAATCGCCAGAACCTTGGATTTTGTCAACCTATGCGCTACGACTTCCATGGC	660
QY	752	TCCTGGAGGAGGCTACACTGGAGAGACAGAGCCCTCTTACAAATACCCGACTGCACACCGGC	811
Db	661	TCCTTGGGGAAGGTCACGGGACATACAGAGCCCTCTTACAAGAGGAAGAAGAGTGGT	720
QY	812	AGCAACGCTTACCTCAATGTGGATTATGTCATGAACACTACTGGAAGGACAAATGAGGACCA	871
Db	721	GCAGCAGCCAGCCTCAACGTGGATGCTGCTGTCCAAAGTGGCTGCAGAAGGGGAGCCCT	780
QY	872	GCTGAGAAGCTCATGTTGGATTCCCTTACCTATGGACACAACTTCATCTGTAGCAACCC	931
Db	781	GCCAGCAAGCTGATCCTTGGCATGCCATCCTACGGAGCTCTCTACACTGGCCTCCTCA	840
QY	932	TCCAAACACTGGAATTTGGTGGCCCCACCTCTGGTGTGGTCTGCTGGGCCCTATGGCAAG	991
Db	841	TCAGACACCAGAGTGGGGGGCCCCAGGCCACAGGGTCTGGCCTCCAGGCGCCTTTCACCAAG	900
QY	992	GAGTCTGGGATCTGGGCTTACTAGAGATCTGTACCTTCTTGAAAATGAGGCACTCAG	1051
Db	901	GAGGGGGGATGCTGGCCTACTATGAAGTCTGCTCCT-----GGAGGGGGGCCACCAAA	954
QY	1052	GGATGGGATGCCCTCAGGAAGTGCCTTATGCTATCAGGGCAATGTGGTGGCTAT	1111
Db	955	CAGAGAAATCAGGATCAGAAGTGGCCTACATCTTCCGGGACAAACAGTGGGCTTT	1014
QY	1112	GACAAACATCAAGAGCTTCGATATTAAGGCTCAATGGCTTAAGCACAACAAATTTGAGGC	1171
Db	1015	GATGATGTGGAGAGCTTCAAAACCAAGGTACGCTATCTGAAGCAGAAGGACTGGCGGG	1074
QY	1172	GCAATGCTTGGCCATTGATCTGGATGACTTTCACCTGGCACTTCTTGCAACAGGGCAAG	1231
Db	1075	GCAATGCTTGGCACTGGACTTATAGATGACTTTTCCGGCTTCTCTCGCAACAGGGCCGA	1134
QY	1232	TTTCCCCTAATCTCCACCTTGAAGAAGGCCCTCGCCCTGCAGAGTGCAAAGTTCACGGCT	1291
Db	1135	TACCCCTCATCAGACGCTACGGCAGGAACCTGAGTCTTCCATACTTGGCTTTCAGGCACC	1194


```
Db 924 CCT-----GGAAGGGGGCCACCAACAGAGAAATCCAGGATCAGAAAGGTGCCCTACATCT 977
Qy 1086 ATCAGGGCAATGTCTGGTTGGCTATGACAACTCAAGAGCTTCGATATTAAGGCTCAAT 1145
Db 978 TCCGGACACACAGTGGTGGCTTTGATGATGAGAGAGCTTCAAAACCAAGGTACGCT 1037
Qy 1146 GGCTTAAGCAACAATTTGGAGGGCCCATGTCTCTGGGCCATTGATCTGGATGACTTCA 1205
Db 1038 ATCTGAAGCAAGAGGACTGGCGGGCCATGTCTGGGCACCTGGACTTAGACTTTG 1097
Qy 1206 CTGGCACTTCTGCAACAGGGCAAGTTTCCCTTAATCTCCACCTGAAGAAGCCCTCG 1265
Db 1098 CCGGCTTCTCTGCAACAGGGCCGATACCCCTCATCCAGAGCTACGGCAGGAACTGA 1157
Qy 1266 GCCTGCAGAGTGAAGTTGCACGGCTCCAGCTCAGCCCAATGAGCCCAATTAAGTCTGCTC 1325
Db 1158 GTCTTCATATCTGCTTCAGGCACCCAG---AGCTTGAAGTTCCAAACAGGTGAGC 1214
Qy 1326 CCAGTGGCAGGGGAACGGAGCGGGAGTAGCAGCTCTGAGGAGCTCGGGAGGCAAGTG 1385
Db 1215 CCTCTGAACCTGAGCA-----TGGCCCCAGCCCTGGACAAGACA 1253
Qy 1386 GATTCCTGTCTGTCAGAGCAACGGCTCTACCCCGTGGCAATAACAGAAATCCCTTCT 1445
Db 1254 CGTCTCCAGGGCAAGCTGTAGGGCTCTATCCCAATCTCGGGAACGGTCCAGCTTCT 1313
Qy 1446 GGCACCTGCGTGAATGAGTACCTTACCAGCAGAACTGCCAGGCGGGCTTGTCTTCGACA 1505
Db 1314 ACAGCTGTGACGGGGCGGCTGTCAGCAAGAGCTGCCGACAGGCTGTGTTCAGCA 1373
Qy 1506 CCAGCTGTGATGCTGCAACTGG 1528
Db 1374 ACTCTGCAAAATGCTGCACCTGG 1396
```

RESULT 14

AAZ21847
ID AAZ21847 standard; DNA; 1636 BP.

XX AAZ21847;

XX 10-DEC-1999 (first entry)

XX MO-218 clone of human Chitinase, with noncoding 5'/3' regions.

XX chitin; fungal infection; immunocompromised; AIDS; chemotherapy;

XX organ transplant; parasite; chitin-binding; allele; vector;

XX truncated protein; ds.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 2..1402

XX /tag= a

XX /product= Human_Chitinase

XX 2..65

XX /tag= b

XX /product= Signal_peptide

XX 66..1402

XX /tag= c

XX /product= Mature_protein

XX W09946390-A1.

XX 16-SEP-1999.

XX 12-MAR-1999; 99WO-US05343.

XX 12-MAR-1998; 98US-0039198.

XX (ICOS-) ICOS CORP.

XX Gray PW, Tjoelker LW;

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

DR

XX

PT

XX

PS

XX

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

Query Match

Best Local Similarity

Matches 853; Conservative

29.68; Score 481; DB 20; Length 1636;

60.8%; Pred. No. 6.6e-126;

Mismatches 520; Indels 30; Gaps 3;

Qy 126 CAGGTCTTGTCTATATCAATTTGCAGCTCGCTCTGCTACCACTGACATGCTACT 185

Db 24 CAGGTTTCATGGTCTCTGCTATGATCCCATGGGGCTCTGCTGCANAACCTGGTCTACT 83

Qy 186 TCACCAACTGGGGCCAGTACCGGCCAGGCGCTGGGGGGCTTCATCGCTGCAACATCGACC 245

Db 84 TCACCAACTGGGGCCAGTACAGAGAGGGGGAGGCTCGCTTCTCTCCCAAGGACTTGGACC 143

Qy 246 CCGTCTCTCTACCACTGATCTAGCCCTTTCTGGGAGGAGAGCAACAGCATCACCA 305

Db 144 CCAGCCCTTTCACCCACCTCATCTAGCCCTTCGCTGGCATGACCAACCACTGAGCA 203

Qy 306 CCATCGAATGGAACGATGTGACTCTCTACCAAGCTTTCAATGGCCTGAAATAAGAACCA 365

Db 204 CCATGAGTGAATGACGAGACTCTCTACCAAGGATTCATGCGCTGAAGAAGATGAATC 263

Qy 366 GCAGCTGAAACTCTCTCGGCCATTTGGAGGCTGGAACTTCGGGAGTCCCTTTCAGTG 425

Db 264 CCAAGCTGAAGACCCCTGTTAGCCATCGGAGCTGGAATTTCCGCACTCAGAAGTTCACAG 323

Qy 426 CCATGTTTCTACTCTCTGAGAACCGCCAGACTTTTCATCCTCAGTCATCAATTCCTGC 485

Db 324 ATATGGTAGCCAGGGCCAAACCCGTCAGACCTTTGTCAACTCGGCCATCAGGTTTCTGC 383

Qy 486 GCCAGTATGAGTTTACGGGCTGGACTTTGACTGGGAGTACCTTGGCTCTCTGGGAGCC 545

Db 384 GCAATACAGCTTTGACGGGCTTTCAGCTTCTGACTGGGAGTACCCAGGAAGCCAGGGAGCC 443

Qy 546 CTCTCAGGACAGCATCTCTTCACTGTCTGTGTGAGGAAATTCGCTGAAGCTTTTGAGC 605

Db 444 CTGCGCTAGACAAGAGCGGCTTCACAACCCCTGTGTACAGGACTTTGGCCAATTCCTTCAGC 503

Qy 606 AGGAGGCCAAGCAGATCAACAAGCCCGGCTGTGTCTACTGCTGCTGAGTAGCTGCTGGCA 665

Db 504 AGGAAGCCCAAGACCTCAGGGAAGGAGACGCCCTTCTCTGAGTGCAGCGGTTCCAGCTGGC 563

WPI; 1999-551417/46.

P-PSDB; AAY42425.

Novel chitin-binding fragments of human chitinase used to treat fungal infections in animals

Example 1; Page 55-57; 83pp; English.

This is the nucleotide sequence of an allelic form of the human chitinase enzyme, which is capable of degrading Chitin (a linear homo polymer of beta-1,4-linked N-acetylglucosamine residues). Chitinase fragments can be used to screen for proteins or other molecules that specifically bind to the chitin-binding domain of human chitinase or that modulate its activity. These compounds are useful for immunization, as well as for purifying chitinase, as well as for detection and quantification of chitinase. Polynucleotide fragments of the invention are useful as a source of probes and primers, and to express the proteins recombinantly. The chitinase fragments, when conjugated to antifungal compounds, are used to treat animals, especially humans, infected with chitin-containing parasites such as fungi. Fungal infection treated include candidiasis, aspergillosis, coccidioidomycosis, blastomycosis, paracoccidioidomycosis, mucormycosis, histoplasmosis, cryptococcosis, chromoblastomycosis, Chitin can be degraded by the enzyme chitinase. Use of whole chitinase protein for treating infections, especially fungal infections, is problematic. In view of the increasing incidents of life-threatening fungal infection in e.g. immunocompromised individuals, there exists a need for identifying new compounds for treating fungal infection. The chitin-binding fragments of the present invention provide this need.

Sequence 1636 BP; 361 A; 491 C; 440 G; 344 T; 0 other;

```
QY 666 TCTCAATATCCAGTCTGGCTATGAGATCCCAACTGTACAGTACCTGGACTACATCC 725
    || || || || || || || || || || || || || || || || || || || ||
Db 564 AGACCTATCTGGATCTGTAGAGTGGAGCAAAATCGCCAGAACCTGGATTTGTCA 623

QY 726 ATGTGATGACCTACGACCTCCATGGCTCTGTGGAGGGGTACACTGGAGAGAACAGCCCC 785
    || || || || || || || || || || || || || || || || || || || ||
Db 624 ACCTTATGGCTTACGACTTCCATGGCTCTGTGGAGAAAGTCCAGGGACATACAGCCCC 683

QY 786 TCTACAAATACCGACTGACACCGGACGCAAGCCCTACCTCAATGTGGATTATGTCATGA 845
    || || || || || || || || || || || || || || || || || || || ||
Db 684 TCTACAAAGGCAAGAAGAGTGTGTGACAGACGAGCTCAACGTGGATGTCTGTCTGC 743

QY 846 ACTACTGGAAGCAATGAGCACCAGCTGAGAAGCTCATCTGTGGATTCCCTTACCTATG 905
    || || || || || || || || || || || || || || || || || || || ||
Db 744 AACAGTGGCTGAGAGGGGACCCCTGCCAGCAGCTGATCTTGGCATGCCCTACCTAGC 803

QY 906 GACACAACTTATCTGTGAGCAACCCCTCCAACTGGAATGTGTGCCCCCACTCTGTGGTG 965
    || || || || || || || || || || || || || || || || || || || ||
Db 804 GACGCTCTTCACTAGCTGGCTCTCATCAGACACAGAGTGGGGGCCCCAGCAGGCT 863

QY 966 CTGTCTCTGCTGGCCCTATGCCAAGGCTCTGGGATCTGGGCTTACTACGAGATCTGTA 1025
    || || || || || || || || || || || || || || || || || || || ||
Db 864 CTGGCACTCCAGGCCCTTCAACAGGAAGAGGAGTGTCTGGCTTACTATGAGTCTGCT 923

QY 1026 CCTTCTGAAAAATGAGCCACTCAGGGATGGGATGCCCTCAGGAAGTGCCTTATGCT 1085
    || || || || || || || || || || || || || || || || || || || ||
Db 924 CCT-----GGAAGGGGCCACCAACAGAGATCCAGGATCAGAGGTGCCCTACATCT 977

QY 1086 ATCAGGGCAATGTGGGTGGCTATGACAACTCAAGAGCTTCGATATTAAGGCTCAAT 1145
    || || || || || || || || || || || || || || || || || || || ||
Db 978 TCCGGACACACCAAGCTGGGTGGCTTTGATGATGTGGAGAGCTTCAAAACCAAGGTCAGCT 1037

QY 1146 GCGTTAAGCACACAAATTTGAGGCGCCATGCTTGGGCCATTTGATCTGGATGACTTCA 1205
    || || || || || || || || || || || || || || || || || || || ||
Db 1038 ATCTGAAGCAGAAAGGACTGGGGGGGCGCATGCTTGGGCACTGGGACTTAGATGACTTGG 1097

QY 1206 CTGGCACTTTCTGCAACCAAGGCAAGTTTCCCTTAATCTCCACCTGAAGAGGCCCTCG 1265
    || || || || || || || || || || || || || || || || || || || ||
Db 1098 CCGGCTTCTCTGCAACCAAGGCGCGATACCCCTCATCCAGAGCTACGCGAGGAACTGA 1157

QY 1266 GCGTCGAGAGTCAAGTTGCACGGCTCCAGCTCAGCCCAATGAGCCCAATTAATGTGCTC 1325
    || || || || || || || || || || || || || || || || || || || ||
Db 1158 GTCTTCCAPACTTGTGCTCAGGACCCCAAG---AGCTTGAAGTTTCCAAAACCAAGGTCAGC 1214

QY 1326 CCAAGTGGCAGCGGACGCGGACGGAGTACAGCTCTGGAGGCACTCGGAGGAGCAGTG 1385
    || || || || || || || || || || || || || || || || || || || ||
Db 1215 CCTCTGAACCTGAGCA-----TGGCCCCAGCCCTGGACAAAGCA 1253

QY 1386 GATTTCTGTCTCTCAGAGCAACGGCTCTACCCGTTGGCAAAATACAGAAATGCTTCT 1445
    || || || || || || || || || || || || || || || || || || || ||
Db 1254 CGTCTGCCAGGCAAGCTGATGGCTCTATCCCAATCTCGGGAACGGTCCAGCTTCT 1313

QY 1446 GGCATGCTGTAATGAGTACGTAACAGCAAGTGCAGCGCGGGCTTGTCTTGGACA 1505
    || || || || || || || || || || || || || || || || || || || ||
Db 1314 ACAGCTGTGACAGCGGGGCTGTCTCCAGAAAGTGCAGAGGCTGTGGTGTTCAGCA 1373

QY 1506 CCAGCTGTGATGCTGCAACTGS 1528
    || || || || || || || || || || || || || || || || || || || ||
Db 1374 ACTCTGCAATGCTGCACCTGG 1396
```

RESULT 15

```
AAD03759
ID AAD03759 standard; cDNA; 1636 BP.
XX
AC AAD03759;
XX
XX
DT 19-JUN-2001 (first entry)
XX
DE Human chitinase cDNA from clone pMO-218.
XX
KW Human; antifungal; chitinase; immunoglobulin; Ig; therapy;
```

```
KW fungal infection; candidiasis; aspergillosis; coccidioidomycosis;
KW blastomycosis; paracoccidioidomycosis; histoplasmosis; cryptococcosis;
KW chromoblastomycosis; sporotrichosis; mucormycosis; dermatophytoses;
KW clone pMO-218; ss.
XX
XX Homo sapiens.
XX Key Location/Qualifiers
XX CDS 2..1402
XX FT /tag= a
XX FT /product= "Human chitinase from clone pMO-218"
XX FT 2..64
XX FT sig_peptide
XX FT /tag= b
XX FT 65..1399
XX FT mat_peptide
XX FT /tag= c
XX FT /product= "Human mature chitinase from clone pMO-218"
XX
XX WO200123430-A2.
XX
XX 05-APR-2001.
XX
XX 28-SEP-2000; 2000WO-US26960.
XX
XX 30-SEP-1999; 99US-0409918.
XX
XX (ICOS-) ICOS CORP.
XX
XX PA Allison DS, Dietsch GN, Gray PW, Shaw KD, Steiner BH;
XX
XX PI
XX
XX WPI; 2001-266141/27.
XX
XX P-PSDB; AAE00432.
XX
XX Novel chitinase immunoglobulin fusion product, useful for treating
XX fungal infections and reducing the amount of a non-chitinase antifungal
XX agent needed for the treatment.
XX
XX Claim 2; Page 29-31; 39pp; English.
XX
XX The present invention relates to a chitinase immunoglobulin (Ig) fusion
XX product, comprising a human chitinase fused to at least a portion of an
XX immunoglobulin chain. The fusion product is useful for treating fungal
XX infections (mycoses) such as candidiasis, aspergillosis, blastomycosis,
XX coccidioidomycosis, paracoccidioidomycosis, histoplasmosis, mucormycosis,
XX cryptococcosis, chromoblastomycosis, sporotrichosis and dermatophytoses.
XX The fusion protein is useful for reducing the amount of non-chitinase
XX antifungal agent needed to exert an antifungal activity. The fusion
XX protein is also useful for preparing a medicament for the prophylactic
XX or therapeutic treatment of fungal infections. Chitinase immunoglobulin
XX fusion product has unexpectedly improved serum half-life and formulation
XX properties. The present sequence is human chitinase cDNA from clone
XX pMO-218. Chitinase enzyme degrades chitin which is a homopolymer of
XX beta-(1,4)-linked N-acetylglucosamine residues.
XX
XX Sequence 1636 BP; 361 A; 491 C; 440 G; 344 T; 0 other;
```

```
Query Match 29.6%; Score 481; DB 22; Length 1636;
Best Local Similarity 60.8%; Pred. No. 6,6e-126;
Matches 853; Conservative 0; Mismatches 520; Indels 30; Gaps 3;

QY 126 CAGGTCTTGTCTTACTTACTGAATTTGACGTCTGGCTCTGCTTACCAAGTGTACT 185
    || || || || || || || || || || || || || || || || || || || ||
Db 24 CAGGTTTCTATGCTCTGCTGATGATCCCATGGGCTCTGCTCAAAACTGGTCTGCTACT 83

QY 186 TCACCAACTGGGCCCCAGTACCGGCCAGGCTGGGGGCTTCATGCTGTGACATCGACC 245
    || || || || || || || || || || || || || || || || || || || ||
Db 84 TCACCAACTGGGCCCCAGTACAGAGGGGGAGGCTGCTTCTCTGCCAAGGACTTGGACC 143

QY 246 CCGTCTCTGTACCCACCTGATCTAGCCTTTGCTGGGAGGAGAGACACGAGATCACC 305
    || || || || || || || || || || || || || || || || || || || ||
Db 144 CCAGCCTTTGACCCCACTCATCTAGCCTTGGCATGACCAACCAAGGCTGAGCA 203

QY 306 CCATGGAATGGAACGATGTGACTCTTACCAAGCTTTCAATGCTTGAAAAATAAGAAC 365
    || || || || || || || || || || || || || || || || || || || ||
```

Db 204 CCACTGAGTGAATGACGAGACTCTCTACAGGAGTTCAATGGCCTGAAGAAGATGAATC 263
Qy 366 GCCAGCTGAATACTCTCTGCGCATTTGGAGCTTGAACCTTCGGGACTGCCCTTTCACTG 425
Db 264 CCAAGCTGAAGACCCCTGTAGCCATCGGAGCTGGAATTTGGGCACTCAGAGTTCCACAG 323
Qy 426 CCATGGTTTACCTCCCTGAGAACCGCCAGACTTTTCACTACCTCAGTCAATCAATTTCTGC 485
Db 324 ATATGGTAGCCACCGCCCAACAACCGTCAGACCTTTGTCAACTCGGCCATCAGGTTTCTGC 383
Qy 486 GCCAGTATGAGTTTGAAGGCTGACCTTGGAGTACCTGAGTACCTGCTGCTGGAGCC 545
Db 384 GCAATACAGCTTTGACGGCTTGAACCTTGAAGTACCCAGGAAGCGGAGGCC 443
Qy 546 CTCCTCAGGCAAGCACTCTTCACTGTCTGTCGAGGAATGCTGAAGCTTTTGTAGC 605
Db 444 CTGCGTAGACAAGAGGCTTCAACACCTTGGTACAGGACTTGGCCAAATGCCCTTCCAGC 503
Qy 606 AGGAGGCCAAGCAGATCAACAAGCCCGAGCTGATGTCACGTGTCAGTACCTGCTGSCA 665
Db 504 AGGAAGCCAGACCTCAGGAAGGAAGCGCTTCTTCTGAGTGCAGCGGTTCCAGCTGGGC 563
Qy 666 TCTCCAATATCCAGTCTGGCTATGAGATCCCGCAACTGTACAGTACCTGAGCTACATCC 725
Db 564 AGCCTATGAGGATGCTGGATGAGGATGAGGAGGACAAATCGCCCAAGACCTGGATTTGTCA 623
Qy 726 ATGTCATGACCTAGCAGCTCCATGGCTCTGGGAGGCTACACTGGAGAGAACAGCCGCC 785
Db 624 ACCTTATGGCTAGGACTTCCATGGCTTGGGAGAGGTCACGGGACATTAACAGCCGCC 683
Qy 786 TCTACAATATCCGAGCTGACACGGGAGCAACGCTTACCTCAATGTGGATTATGTCATGA 845
Db 684 TCTACAAGAGCAAGAGAGAGTGGTGCAGCAGCCGCTCAAGCTGGATGCTGCTGTC 743
Qy 846 ACTACTGGAAGGCAATGGAGCACCAGCTGAGAGCTCATCGTTGGATTCCCTACCTATG 905
Db 744 AACAGTGGCTGAGAGGGGACCCCTGCCAGCAGCTGATCCTTGGCATGCCCTACCTACG 803
Qy 906 GACAACTTTCATCCTGAGCAACCCCTCCAACATGGAATGGTGGCCGCCACCTCTGCTG 965
Db 804 GAGCTCTTTCACACTGGCTCTCATCAGACACAGAGTGGGGGCCGCCAGCCACAGGT 863
Qy 966 CTGCTCTGCTGGCCCTATGCCAAGAGCTGCGGATCTGGGCTTACTAGAGATCTGTA 1025
Db 864 CTGSCACTCCAGGCCCTTCCCAAGAGAGGAGTGTGGCTTACTATGAAGTCTGCT 923
Qy 1026 CTTCTCTGAAATGGAGCAGCTCAGGATGGGATGCCCTCAGGAAGTGCCTTATGCT 1085
Db 924 CCT-----GGAAGGGGCCCAACACAGAGATCCAGGATCAGAAGGTGCCCTACATCT 977
Qy 1086 ATCAGGCAATGTGTGGTTGGCTATGACAACTCAAGAGCTTCGATATTAAAGGCTCAAT 1145
Db 978 TCCGGGACACACCGAGTGGGCTTTGATGATGAGAGCTTCAAAACCAAGGTCACT 1037
Qy 1146 GGCTTAAGCACACAAATTTGGAGGCCCATGCTGGGCCATGATCTGGATGACTTCA 1205
Db 1038 ATCTGAAGCAGAGAGGACTGGGGGGCCATGCTGGGCACTGGACTTAGATGACTTG 1097
Qy 1206 CTGSCACTTTCGCAACAGGCAAGTTTCCCTTATCTCCACCTCAAGAGGCCCTCG 1265
Db 1098 CCGGCTTCTCCTGCAACAGGGGCGGATACCCCTCATCCAGAGCTACGGCAGGAATGA 1157
Qy 1266 GCCTGAGAGTGAAGTTGACCGGCTCCAGCTCAGCCCATTTGAGCCCAATAACTGCTGCTC 1325
Db 1158 GTCTTCACTATTGCTTACAGGACCCAG---AGCTTGAAGTTCCAAACACAGGTACGC 1214
Qy 1326 CAGTGGCAGGGAGAGGGAGGAGTAGCAGCTCTGAGGAGCTCGGGAGGAGTG 1385
Db 1215 CCTCTGAACCTGAGCA-----TGGCCCCAGCCCTGGAGAGACA 1253
Qy 1386 GATTCGTGCTGTCAGAGCAAGGCTCTACCCCTGGCAATAACAGAAATGCCCTTCT 1445
Db 1254 CGTCTGCCAGGGCAAGCTGATGGGCTCTATCCCAATCCTCGGGAACGCTCCAGCTTCT 1313

Qy 1446 GGCACCTGCTGAATGGAGTCACTACAGCAGAACTGCCAGCGGGCTTGTCTTCGACA 1505
Db 1314 ACAGCTGTGCAGCGGGGCTGTCCAGCAAAAGTCCCCACAGGCTGTGTTCAGCA 1373
Qy 1506 CCAGCTGTGATGTGCTGCAACTGG 1528
Db 1374 ACTCCTGCAATGCTGCACCTGG 1396

Search completed: July 3, 2003, 05:03:18
Job time : 377.46 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2003, 04:38:16 ; Search time 78.4127 Seconds
(without alignments)
6355.472 Million cell updates/sec

Title: US-10-004-219B-2
Perfect score: 1625
Sequence: 1 gctttcagtcgtgtgtga.....ctgcaataaaatcagcagtc 1625

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*

- 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/PCRUS_COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	483	29.7	1656	US-09-039-198A-3	Sequence 3, Appl
2	483	29.7	1656	US-08-877-599-3	Sequence 3, Appl
3	483	29.7	1656	US-09-267-574-3	Sequence 3, Appl
4	481.4	29.6	1643	US-08-486-839-3	Sequence 3, Appl
5	481.4	29.6	1643	US-09-151-011-3	Sequence 3, Appl
6	481.4	29.6	1643	US-09-343-623-3	Sequence 3, Appl
7	481	29.6	1636	US-09-039-198A-1	Sequence 1, Appl
8	481	29.6	1636	US-08-877-599-1	Sequence 1, Appl
9	481	29.6	1636	US-09-267-574-1	Sequence 1, Appl
10	477.2	29.4	1713	US-08-486-839-5	Sequence 5, Appl
11	477.2	29.4	1713	US-09-151-011-5	Sequence 5, Appl
12	477.2	29.4	1713	US-09-343-623-5	Sequence 5, Appl
13	369	22.7	1433	US-08-694-915-1	Sequence 1, Appl
14	354.8	21.8	1526	US-08-694-915-3	Sequence 3, Appl
15	339.8	20.9	1681	PCT-US94-07754-4	Sequence 4, Appl
16	338.2	20.8	1681	US-08-581-527-4	Sequence 4, Appl
17	156	9.6	2452	US-08-524-051-1	Sequence 1, Appl
18	127.2	7.8	1677	US-09-545-814-13	Sequence 13, Appl
19	127.2	7.8	1677	US-09-545-814-15	Sequence 15, Appl
20	127.2	7.8	1749	US-09-545-814-4	Sequence 4, Appl
21	127.2	7.8	1749	US-09-545-814-6	Sequence 6, Appl
22	127.2	7.8	1919	US-09-545-814-31	Sequence 31, Appl
23	127.2	7.8	1919	US-09-545-814-33	Sequence 33, Appl
24	127.2	7.8	2610	US-09-545-814-1	Sequence 1, Appl
25	127.2	7.8	2610	US-09-545-814-3	Sequence 3, Appl
26	120.6	7.4	1478	US-09-545-814-28	Sequence 28, Appl
27	120.6	7.4	1478	US-09-545-814-30	Sequence 30, Appl

Sequence 20, Appl
Sequence 22, Appl
Sequence 17, Appl
Sequence 19, Appl
Sequence 14, Appl
Sequence 16, Appl
Sequence 40, Appl
Sequence 42, Appl
Sequence 37, Appl
Sequence 39, Appl
Sequence 34, Appl
Sequence 36, Appl
Sequence 6, Appl
Sequence 10, Appl
Sequence 14, Appl
Sequence 13, Appl
Sequence 9, Appl
Sequence 11, Appl

ALIGNMENTS

RESULT 1
US-09-039-198A-3
; Sequence 3, Application US/09039198A
; Patent No. 6200951
; GENERAL INFORMATION:
; APPLICANT: Gray, Patrick W.
; TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive/6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/039,198A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 27866/34391
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1656 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 27..1424
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 90..1424
; US-09-039-198A-3

Query Match 29.7%; Score 483; DB 4; Length 1656;
Best Local Similarity 60.2%; Pred. No. 4.6e-132;
Matches 870; Conservative 0; Mismatches 545; Indels 30; Gaps 3;

144 TGAATTTGACGCTGGCTCTGCTTACAGCTGACATGCTACTTCCACCACTGGGCCAGT 203
145 TGAATTTGACGCTGGCTCTGCTTACAGCTGACATGCTACTTCCACCACTGGGCCAGT 126
146 TGAATTTGACGCTGGCTCTGCTTACAGCTGACATGCTACTTCCACCACTGGGCCAGT 126
204 ACCGGCCAGGCTGGGCGCTTATGCTGCTGACACATGACCCCTGCTCTGTACCCACC 263
127 ACAGACAGGGGGGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 186
264 TGATCTACGCTTGTGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 323
187 TCATCTACGCTTGTGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 246
324 TGATCTCTACGCTTGTGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 383
247 AGATCTCTACGCTTGTGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 306
384 TGCCATTTGAGGCTGGAATTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 443
307 TAGCCATCGGAGGCTGGAATTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 366
444 AGAACCCGACGCTTGTGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 503
367 ACACCGCTACGCTTGTGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 426
504 GGCTGACTTGTGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 563
427 GCCTTGACCTTGTGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 486
564 TCCTCACTGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 623
487 GCTTCACAGCTTGTGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 546
624 ACAAGCCGACGCTTGTGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 683
547 GGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 606
684 GCTATGAGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 743
607 GATACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 666
744 TCATGCTGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 803
667 TCATGCTGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 726
804 ACACCGCTGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 863
727 AGAGTGTGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 786
864 GAGCAGG 923
787 GGACCCCTGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 846
924 GCAACCCCTGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 983
847 CTTCTCTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 906
984 ATGCCAAGG 1043
907 TCACCAAGG 960
1044 CCACTCAGG 1103
961 CCACCAAGG 1020
1104 TTGCTATGACACATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1163
1021 TGGCTTTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1080
1164 TTGAGGCGGAGG 1223
1081 TGGGCGGAGG 1140
1224 AGGCGAAGTTTCCCTTAACTCCACCTGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1283

1141 AGGCCGATACCCCTCATCCAGACGCTAGCGCAGGAACCTGAGTCTTCCATACCTTGGCTT 1200
1284 GCAGGCTCCAGCTCAGCCCATGAGCAATACTGCTGCTCCAGTGGCAGCGGAGACG 1343
1201 CAGCAGCCCGC---AGCTTGAAGTTCCAAACAGGTCAGCCCTCTGAACCTGAGCA-- 1255
1344 GGAGCGGAGTAGCAGCTCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1403
1256 -----TGGCCCCAGCCCTGGCAGACACGCTTCTGCCAGGCGAAG 1296
1404 CCAACGCGCTTACCCGCTGGCAATAAAGAAATGCTTCTTGGCAGTGGGTGAATGGAG 1463
1297 CTGATGGCTCTATCCCAATCTCGGAGCGGTCAGCTTCTACAGCTGTGCAGCGGGC 1356
1464 TCAGTACCAGCAACTGCCAGGCGCGGCTTCTTCCAGACACGCTGTGATTCGTCGCA 1523
1357 GGCTGTTCAGCAAGCTGCCGAGCGCTGCTTTCAGCAACTCTGCAAAATGCTGCA 1416
1524 ACTGG 1528
1417 CTGG 1421

RESULT 4
US-08-486-839-3
; Sequence 3, Application US/08486839
; Patent No. 5928928
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: A human chitinase, its recombinant
; TITLE OF INVENTION: production, its use for decomposing chitin, its use
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11758
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; SOFTWARE:
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,839
; FILING DATE: 07 - June - 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Baron, Ronald J.
; REGISTRATION NUMBER: 29,281
; REFERENCE/DOCKET NUMBER: 294-26
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1643 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; HYPOTHEetical: NO
; US-08-486-839-3

Query Match 29.6%; Score 481.4; DB 2; Length 1643;
Best Local Similarity 60.3%; Pred. No. 1.3e-131;
Matches 866; Conservative 0; Mismatches 541; Indels 30; Gaps 3;
92 CTGACTGCAACCATGACAAAGCTTATCTCTCTCAGAGTCTTGTCTTATCTGAAATTTG 151

Db	1	CTGAGCTGCATCATGCTGGCGTCTGTGGCTGGGCAAGGTTTCATGCTCTCTGTGATGATC	60
Qy	152	CAGCTCGGCTCTGCCTTACCAGCTGACATGCTACTTCCACCAACTGGGCCAGTACCGGCCA	211
Db	61	CCATGGGCTCTGCTCCAAAAGTGGTCTGCTACTTCCACCACTGGGCCAGTACACACAG	120
Qy	212	GGCCTGGGGGCTTCATGCTGTACACATCGACCCCTGCTCTGTATACCCACCTGTATCTAC	271
Db	121	GGGAGGCTCGCTTCCTGCGCCAAAGGACTTGGACCCAGCGCTTGCACCCACCTCATCTAC	180
Qy	272	GCCTTTGCTGGGAGGACACAGATACACACCATCGAATGGAACGATGTGACTCTC	331
Db	181	GCCTTCGCTGGCATGACCAACCAACAGCTGAGCACCATCGATGGGAATGACGAGACTCTC	240
Qy	332	TACCAAGCTTTCAATGGCCTGAAAAATAAGAACAGCCAGCTGAATACTCTCTGGCCATT	391
Db	241	TACCAGAGTTCAATGGCCTGAGAGATGAATCCCAAGCTGAAGACCTGTTAGCCATC	300
Qy	392	GGAGCTGGAACTTTGGGAGCTGCCCCCTTTCACATGCCATGGTTTCTACTCCTGAGAACCG	451
Db	301	GGAGGCTGGAATTTTCGCGACTCAGAAAGTTACAGATATGCTAGCCACGGCCAAACAACCT	360
Qy	452	CAGACTTTCATCACTCAGTCATCAATTCCTCGCGCCAGTATGATGATTTGACGGCTGGAC	511
Db	361	CAGACCTTTGTCAACTGGGCACTCAGGTTTCTGCGCAATACAGCTTTTACGGCCCTTGAC	420
Qy	512	TTTGACTGGAGTACCTTGCTCTCGTGGGAGCCCTCCTCAGGACAAGCATCTCTTCACT	571
Db	421	CTTGACTGGAGTTACCCAGGAGCCAGGGAGGCCCTGCCGTAGACAAGGAGCGCTTCACA	480
Qy	572	GTCCCTGGTGAGGAATTCGGTGAAGCTTTTGGACGAGGCGCCAAAGCAGATCAACAAGCC	631
Db	481	ACCTGCTACAGACTTGGCCAAATGCTTCCAGCAGGAAGCCACAGACCTCAGGGAAGGA	540
Qy	632	AGGCTGATGCTACTGCTGCAGTAGCTGCTGGCATCTCCAATATCCAGTGTGCTATGAG	691
Db	541	CGCCTTCTCTGAGTGCAGGGTTCCAGCTGGGCGACCTATGTGGATGCTGGATACGAG	600
Qy	692	ATCCCCAACTGCAGACTACCTGGACTACATCCATGTCTATGACCTACGACCTCCCATGGC	751
Db	601	GTGACAAATCGCCAGAACCTGGATTTGTCAACCTTATGGCCTACGACTTCCATGGC	660
Qy	752	TCTCTGGAGGGCTACACTGGAGAGAACAGCCCCCTCTACAATATCCCGACTGACACCGGC	811
Db	661	TCTTGGGAGAAGTCAACGGACATACAGCCCCCTCTACAAGGAGCAAGAGAGATGGT	720
Qy	812	AGCAAGCCCTACCTCAATGTGGATTATGTATGAATCTACTGGAAGGACATGAGGACCA	871
Db	721	GCAGACCGACGCTCAACGTGGATGTGCTGTGCAACAGTGGCTGCAGAAGGGGACCCCT	780
Qy	872	GCTGGAAGCTCATCTGTGGATTCCCTACTATGGACACAACCTTCATCTCTGACCAACCC	931
Db	781	GCACGACGTGATCTTGGCATGCCCTACCTACGGAGCTCTCTTACATCTGGCCTCTCTCA	840
Qy	932	TCCAACATGGAAATGGTGGCCCCACCTCTGTGTGCTGGTCTGCTGGGCCCTATGCCAAG	991
Db	841	TCAGACACCAAGTGGGGGCCACGACACAGGCTCTGGCATCCAGGCCCTTCAACCTTCA	900
Qy	992	GAGTCTGGGATCTGGGCTTACTAGGATCTGTACCTTCTGAAAAATGAGGCCACTCAG	1051
Db	901	GAAGGAGGGATGCTGGCCTACTATGAAGTCTGTCTCT-----GGAAGGGGGCGCACAA	954
Qy	1052	GGATGGATGCCCTCAGGAAGTGCCTTATGCTCTATCAGGGCAATGTGTGGTGTGCTAT	1111
Db	955	CAGAGAAATCAGATCAGAAGGTGCCCTACATCTCCGGGACAACCATGTTGGTGGCTTT	1014
Qy	1112	GACAACATCAAGCTTTCGATATTAAGGCTTCAATGGCTTAAGCACACAATAATTGGAGGC	1171
Db	1015	GATGATGTGAGAGCTTCAAAACCAAGGTCAGCTATCTGAAGCAGAAGGACTGGCGGG	1074
Qy	1172	GCCATGGTCTGGGCCATTGATCTGGATGACTTCACTGGGCATTTCTTGCACACAGGCGAAG	1231
Db	1075	GCCATGGTCTGGGCATGGACTTATAGATGACTTTCGGGGCTTCTCTCTGCAACAGGCCGA	1134

RESULT-5
US-09-151-011-3
Sequence 3, Application US-09151011

APPLICANT: A Human Chitinase, Its Recombinant
 TITLE OF INVENTION: Production, Its Use For Decomposing Chitin, Its Use in
 TITLE OF INVENTION: Therapy or Prophylaxis Against Infection Diseases.
 NUMBER OF SEQUENCES: 16
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Hoffmann & Baron, LLP
 STREET: 6900 Jericho Turnpike
 CITY: Syosset
 STATE: New York
 COUNTRY: United States of America
 ZIP: 11791
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30

Query Match	29.6%	Score 481.4;	DB 3;	Length 1643;
Best Local Similarity	60.3%	Pred. NO. 1.3e-131;		
Matches 866;	Conservative	0;	Mismatches 541;	Indels
				30;
				Gaps
				3;

QY 92 CTGACTGCAACCATGACAAAACGCTATTCTCTCTCACAGCTCTGTCTCCTTATCTGAATTG 151
|||||
Db 1 CTGAGCTCATCATGGTGGCGGTGTGGCGCTGGGCAGGTTTCATGGTCTGCTGATGATC 60
|||||

[illegible]

Db	1098	CGGCTTCTCCTGGAACCAAGGGCGGATACCCCTCATCCAGACGCTACGCCAGGAAC	TGA 1157	
Qy	1266	GCCTGCAGAGTGCAAAGTTGCACGGCTCCAGCTCAGGCCATATGAGCCCAATAACTGCTGC	TTC 1325	
Db	11158	GTCTTCCATACTTGCCTTCCAGGCA	CCCGAG --- AGCTTGAAGTTCCAAAACCAAGGTCA	G 1214
Qy	1326	CCAGTGGCAGCGGGACGGAGCGGGAGTAGCAGCTCTG	CAGCAGCTCGGGAGGCGAGTG 1385	
Db	1215	CCCTGAACTGAGCA	-----TGCCCCAGCCCTGGACAGACA 1253	
Qy	1386	GATTCTGTGTGTGCAGAGCCACGGCCTTACCCCGTGGCAATAACAGAAATGCCTTCT	1445	
Db	1254	CGTTCTGCCAGGGCAAGCTGATGGGTCTATCCCAATCCTCGGGACGGTCCAGCTTCT	1313	
Qy	1446	GGCACTCGGTGAATGGAGTCACTTACAGCAGAACTGCCAGCGCGGGCTTGCTTCGACA	1505	
Db	1314	ACAGCTGTGACGGGGCGGCTGTTCAGCAAAAGCTGCCCGACAGGCGCTGGTTCAGCA	1373	
Qy	1506	CCAGCTGTGATTTGCTGCAACTGG	1528	
Db	1374	ACTCCTGCAAAATGCTGCACCTGG	1396	

RESULT 8

```

RES001 8
US-08-877-599-1
; Sequence 1, Application US/08877599
; Patent No. 6372212
; GENERAL INFORMATION:
; APPLICANT: Gray, Patrick W.
; TITLE OF INVENTION: Chitinase Materials and Methods
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/877,599
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/663,618
; FILING DATE: 14-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 27866/33994
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1636 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..1399
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 65..1399
US-08-877-599-1

```


DEPT. OF

Query Match	29.6%	Score 481:	DB 4:	Length 1636:
-------------	-------	------------	-------	--------------

Query Match 29.6%; Score 481; DB 4; Length 1636;
Best Local Similarity 60.8%; Pred. No. 1.8e-131;

Qy	126	CAGTCTTGTGCTTTATAC	TGAAATTTGCAGCTCGGCT	TGCCTACCGAGCTGCAT	CTACT	185
Db	24	CAGCTTTCATGGTCCTGCT	GATGATCCCATGGGGCT	CTGCTGCAAACT	TGCTCTACT	83
Qy	186	TCACCAACTGGGCCGATAC	CGGCGCAGGCTTGGGGCG	CTTCATGCCTGACAACAT	TCGACC	245
Db	84	TCACCAACTGGGCCGAGTAC	AGACAGGGGAGGCTCGCT	TCTGCCCAAGSACT	TGGACC	143
Qy	246	CCTGCCCTGTATACCACT	GTATACAGCTTTTGTGGG	AGCGCAACACGAGAT	TCACCA	305
Db	144	CCAGCCTTTGGACCCACCT	CATCTACGCCTTTCGTGG	CATGACCAACCACCGCT	GAGCA	203
Qy	306	CCATCGAATGGAGCAGATG	TGACTCTCTACCAAGCT	TTTCAATGGCGCTGAAAA	ATAAGAACA	365

Db 204 CCACGTAGTGAATGACGAGACTCTCTACAGGAGTTCAATGGCTGAAGAAGATGATC 263
QY GCCAGCTGAAACACTCTCTGCGCATTTGGAGCTGGAACCTTCGGGACTGCCCTTTCACTG 425
Db 264 CCAAGCTGAACACCTGTTAGCCATCGGAGCTGGAATTTCCGCACTCAGAAGTTACAG 323
QY CCATGGTTTCTACTCTGAGAACGCCAGACTTTTCATCACCTCAGTCTCAAAATTCCTGC 485
Db 324 ATATGGTAGCCACGSCCAACAACCGTCAGACCTTTGTCAACTCGGCCATCAGTTTCTGC 383
QY GCCAGTATGATTTTACGGGCTGACCTTTGACTGGGAGTACCCTGGCTCTCTGTTGGAGCC 545
Db 384 GCAATACAGCTTTGACGGCTTACCTTTGACTGGGAGTACCAGGAGCCAGCGGAGCC 443
QY CTCCTCAGGCAAGCATCTCTTCACTGCTCTGCTGTCAGGAAATCGTGAAGCTTTTGAGC 605
Db 444 CTGCGGTAGACAAGGAGGCTTCAACACCTTGTACAGGACTTGGCCATGCCCTTCAGC 503
QY AGGAGGCCAAGCAGATCAACAAGCCAGGCTGATGGTCACTGCTGCACTGCTGCGCA 665
Db 504 AGGAAGCCAGACCTCAGGGAAGGAAGCCCTTCTTCTGAGTGCAGCGGTTCCAGCTGGC 563
QY TCTCAATATCCAGTCTGGCTATGATCCCTCCCACTGTCAGTACCTGACCTACATCC 725
Db 564 AGACTATGTGGATGCTGGATACAGGAGTGGACAATACTGCCGAACCTGGATTTGTCA 623
QY ATGTCATGACCTACGACCTCCATGGCTCTGCGAGGCTACACTGGAGAGAACAGCCGCC 785
Db 624 ACCTATGGCTACGACTTCCATGGCTTTGGAGAGGTCACGGGACATACAGCCGCC 683
QY TCTCAATATCCGACTGACACCGGCGAGAACGCTTACCTCAATGTGGATATGTCANGA 845
Db 684 TCTACAAGAGCAAGAAGAGAGTGGTGCAGCAGCAGCCTCAACGTGGATGCTCTGTC 743
QY ACTAGTGAAGCAATGAGCACCAGCTGAGAGCTCATGTTGGATTCCTTACCTATG 905
Db 744 AACAGTGCCTGCAAGAGGAGCCCTGCCAGAGCTGATCCTTGGCATGCTACCTACG 803
QY GACACAACTTCATCTGAGCAACCCCTCAACACTGGAATGTTGGTCCCACTCTGCTG 965
Db 804 GACGCTCTTCACACTGGCTCTCTCATCAGACACAGAGTGGGGGCCCAACAGGCT 863
QY CTGCTCTGCTGGCCCTATGCAAGAGTCTGGGATCTGGGCTTACTAGAGATCTGTA 1025
Db 864 CTGCACTCCAGGCGCCCTTCCAAAGGAAGAGGATGCTGGCTTACTATGAATCTGCT 923
QY CTTTCTCTGAAATGAGGCACTCAGGATGGGATGCCCTCAGGAAGTGCCTTATGCT 1085
Db 924 CCT-----GGAAGGGCCACCAACAGAGATCCAGGATCAGAAGTGGCCCTACATCT 977
QY ATCAGGCAATGTGTGGTTGGCTATGACAACTCAAGAGCTTCGATATTAAGCTCAAT 1145
Db 978 TCCGGACACCACTGGTGGCTTTGATGATGGAGAGCTTCAAAACCAAGTCACT 1037
QY GGCTTAGCACAACAATTTGGAGGCCCATGCTGGGCAATGATCTGATGACTTCA 1205
Db 1038 ATCTGAAGCAAGAGGACTGGGGGGCCATGCTGGGCACTGGACTTAGATGACTTTG 1097
QY CTGCACTTTCTGCAACCAAGGCAAGTTTCCCTTAATCTCCACCTCAAGAAGGCCCTG 1265
Db 1098 CCGCTTCTCTGCAACAGGCGGATACCCCTTCATCCAGACGCTACGGCAGAACTGA 1157
QY GCCTGCAAGTGAAGTTGACCGGCTCCAGCTCAGGCCATTTAGGCCAATAACTGCTCTC 1325
Db 1158 GTCTTCCATATCTTGGCTTTCAGGACCCAG---AGCTTGAAGTTTCCAAACCAAGTCA 1214
QY CCAGTGCAGCGGGAACGGAGCGGAGTAGCAGCTCTGGAGGAGCTCGGGAGGAGTG 1385
Db 1215 CCTTGAACCTGAGCA-----TGCCCGCCAGCCCTGGACAGACA 1253
QY GATTCTGTGCTGTCAGAGCAAGCCGCTCTACCCGCTGGCAATAACAGAAATGCCCTCT 1445
Db 1254 CGTCTCCAGGCAAGCTGATGGGCTCTATCCCAATCTCTCGGGAACGGTCCAGCTTCT 1313

QY 1446 GGCACTGGTGAATGAGTACAGTACAGCAGAACTGCCAGGCCGGCTTCTCTTCGACA 1505
Db 1314 ACAGCTGTGAGCGGGGGCTGTTCAGCAAAAGCTGCCGACAGGCTGGTGTTCAGCA 1373
QY 1506 CCAGCTGTGATGCTGCAACTGG 1528
Db 1374 ACTCTGCAATGCTGCACCTGG 1396

RESULT 10.

US-08-486-839-5
; Sequence 5, Application US/08486839
; Patent No. 5928928
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: A human chitinase, its recombinant
; production, its use for decomposing chitin, its use
; TITLE OF INVENTION: In therapy or prophylaxis against infection diseases.
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11758
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; SOFTWARE:
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,839
; FILING DATE: 07 - June - 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Baron, Ronald J.
; REGISTRATION NUMBER: 29,281
; REFERENCE/DOCKET NUMBER: 294-26
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1713 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
US-08-486-839-5

Query Match 29.4%; Score 477.2; DB 2; Length 1713;
Best Local Similarity 62.7%; Pred. No. 2.4e-130;
Matches 762; Conservative 0; Mismatches 448; Indels 6; Gaps 1;
QY 92 CTGACTGCAACCATGACAAAGCTTATTCCTCCTCAGAGTCTTCTCTTACTACTGAAATTTG 151
Db 1 CTGAGCTGCATCATGTTGCGGCTCTGCGCTGGCAGGTTTCATGGTCTGCTGATGATC 60
QY 152 CAGCTCGGCTCTGCCCTACCACTGACATCTTCAACCACTGGGCGCCAGTACCGGCCA 211
Db 61 CCATGGGCTCTGCTCCAAACTGGTCTACTTCCAACTGGGCGCCAGTACAGACAG 120
QY 212 GGCTTGGGGGCTTATGCTGCTGACACATCGACCCCTGCTCTGTACCCACCTGATCTAC 271
Db 121 GGGAGGCTGCTCTCTCTCCCAAGGACTTGGACCCCGAGCTTTGCACCCACCTCATCTAC 180
QY 272 GCCTTGTGGGAGGAGAACACAGATCACCACCATCAATGAATGAAGCAATGTGACTCTC 331
Db 181 GCCTTGTGGCTATGACCAACACCACTGAGTGGATGACAGAGACTCTC 240

```
QY 332 TACCAGCTTTCAATGGCTGAAATAAGAACAGCCAGCTGAAAACTCTCGGCCATT 391
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
QY 241 TACCAGAGTTCAATGGCTGAAAGATGAATCCCAAGCTGAAGACCTGTAGCCATC 300
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
QY 392 GGAGCTGGAATTCGGGACTGCCCTTTTCACTGCCATGGTTTCTACTCTCGAGAACCCG 451
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
QY 301 GGAGCTGGAATTCGGGACTCGAAGATTCACAGATATGTTAGCCAGGCCAACACCGT 360
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
QY 452 CAGACTTTCATCACTCAGTCATCAATTCCTGGCCAGTATGAGTTTGGAGGCTGGAC 511
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
QY 361 CAGACTTTCATCACTCAGTCATCAATTCCTGGCCAGTATGAGTTTGGAGGCTGGAC 420
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
QY 512 TTTGACTGGAGTACCTGGCTCTCTGGGGAGCCCTCTCAGGACAGCATCTCTTCACT 571
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
QY 421 CTTGACTGGAGTACCCAGGAGCCAGGAGCCCTGCTGCTAGACAGGAGCGCTTCA 480
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
QY 572 GTCTGTGTGAGGAATTCGTGAGCTTTTGGAGGAGGCGCAAGCAGATCAACAGCCG 631
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
QY 481 ACCCTGTACAGGACTTGGCCAATGCTTCCAGCAGGAAGCCAGACCTCAGGGAAGAA 540
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
QY 632 AGCTGATGCTCACTGCTGAGTACGCTGGCATCTCCATATCCAGTCTGGCTATGAG 691
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
QY 541 CGCTTCTCTGAGTGGAGCGGTTCACGTGGGAGACCTTATGGGATGCTGGATAGAG 600
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
QY 692 ATCCCCAACTGTACAGTACCTGGACTACATCATGCTACGACTACGACCTCATGGC 751
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
QY 601 GTGACAAATCGCCAGAACCTGGATTTTGTCAACCTTATGGCTACGACTTCCATGGC 660
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
QY 752 TCCTGGAGGGCTACACTGGAGAGAACAGCCCTCTTCAAAATACCCGACTGACACCGCC 811
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
QY 661 TCCTGGAGAGGCTCAGCGGACATACAGCCCTCTTCAAGAGGCAAGAGAGTGGT 720
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
QY 812 AGCAAGCCCTACCTCAATGTGGATATGTCATGAATCTTGGAGGAGCAATGGAGCACA 871
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
QY 721 GCAGCAGCAGCTCAACGTGGATGCTGTGCAACAGTGGCTGCAAGAGGAGGACCCCT 780
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
QY 872 GCTGAGAGCTCATCTGTGGATTCCTTACCTATGGACACAACTTCATCTTGAGCAACCCC 931
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
QY 781 GCCAGCAAGCTGATCTTGGATGCTTACCTACCTACGAGGCTTCTTCACTGCGCTCTCA 840
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
QY 932 TCCAACACTGGAATTTGGTGGCCCACTCTGGTGGTGGTCTGCTGGGCGCCCTATGCCAAG 991
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
QY 841 TCAGACACCAAGTGGGGGCCCCAGCCAGGGTCTGGCACTCCAGGCCCTTCCACCAAG 900
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
QY 992 GAGTCTGGGATCTGGCTTACTACGAGATCTGTACCTTCTGAAATATGAGCCACTCAG 1051
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
QY 901 GAAGGAGGATGCTGGCTTACTATGAAGTGTGCTCT- - - - -GGAAGGGGGCCACAAA 954
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
QY 1052 GGATGGATGCCCTCAGGAAGTGCCTTATGCTCTATCAGGGCAATGTGGTGGCTAT 1111
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
QY 955 CAGAGATCCAGGATCAGAGTGGCTTACATCTTCCGGACACCACTGAGTGGGCTTT 1014
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
QY 1112 GACAACATCAAGAGCTTGAATATTAAGGCTCAATGGCTTAAGCACAAATAATTTGAGGC 1171
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
QY 1015 GATGATGTGAGAGCTTCAAAACCAAGGTCAGCTATCTGAAGCAGAGGAGTGGGGGG 1074
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
QY 1172 GCCATGCTTGGGCCATTTGATCTGGATGATCTTCACTGGCACTTCTGCAACCAAGGCAAG 1231
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
QY 1075 GCCATGCTTGGGCACTTGGACTTATGATGACTTTGGCGGCTTCTCTGCAACCAAGGCGCA 1134
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
QY 1232 TTTCCCTTAATCTCCACCTTGAAGAGGCGCTCGGCTGAGAGTGAAGTGTGACGCGCT 1291
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
QY 1135 TACCCCTCATCAGAGCTTACCGCAGGAATCAATGGGTAAGCCCTTAATCTGCTCTCA 1194
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
QY 1292 CCAGCTCAGCCCATTTG 1307
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
QY 1195 CATGTGAGGCCAGGTG 1210
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
```

RESULT 11

US-09-151-011-5

; Sequence 5, Application US/09151011

; Patent No. 6057142

```
GENERAL INFORMATION:
APPLICANT: A Human Chitinase, Its Recombinant
TITLE OF INVENTION: Production, Its Use For Decomposing Chitin, Its Use in
TITLE OF INVENTION: Therapy or Prophylaxis Against Infection Diseases.
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hoffmann & Baron, LLP
STREET: 6900 Jericho Turnpike
CITY: Syosset
STATE: New York
COUNTRY: United States of America
ZIP: 11791
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
SOFTWARE:
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/151,011
FILING DATE: 10 - September - 1998
ATTORNEY/AGENT INFORMATION:
NAME: Morris, Robert C.
REGISTRATION NUMBER: 42,910
REFERENCE/DOCKET NUMBER: 294-32 DIV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 822-3550
TELEFAX: (516) 822-3582
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1713 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
US-09-151-011-5
```

Query Match 29.4%; Score 477.2; DB 3; Length 1713;
Best Local Similarity 62.7%; Pred. No. 2.4e-130;
Matches 762; Conservative 0; Mismatches 448; Indels 6; Gaps 1;

```
QY 92 CTGACTCCAAACCATGACAAAGCTTATCTCTCTACAGGTCTTGTCCTTATATCAATTG 151
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
QY 1 CTGAGCTGCATCATGTGTGGCTTGTGGCGAGGTTTCATGGTCTGCTGATGATC 60
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
QY 152 CAGCTCGGCTCTGCTTACCAGCTGACATGTCTTACCACTGGGCCAGTACCGGCA 211
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
QY 61 CCATGGGCTCTGCTCCAAACTGGTCTGCTACTTACCACTGGGCCAGTACAGAC 120
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
QY 212 GGCTTGGGGCGCTTTCATGCTTGACAAATCGACCCCTGCTCTGTACCCAGTATCTAC 271
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
QY 121 GGGAGGCTCGCTTCTCTGCCAAGGACTTGGACCCAGCCTTTGCACCCACCTCATCTAC 180
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
QY 272 GCCTTCTGGGAGGAGAACAGGATCACCACCATCGAATGGAACGATGTGACTCTC 331
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
QY 181 GCCTTCTGGCATGACCAACCACTGAGTGGATGAGGAGACTCTC 240
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
QY 332 TACCAAGCTTTCAATGGCTTGAATAAAGAACAGCCAGCTGAAACTCTCTGGCCATT 391
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
QY 241 TACCAGAGTTCAATGGCTTGAAGATGAATCCCAAGCTGAGACCTCTGTAGACATC 300
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
QY 392 GGAGCTGGAATTCGGGAGTGGCCCTTTTCACTGCCATGTTTCTACTCTCTGAGAAGCCG 451
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
QY 301 GGAGCTGGAATTTGGGCACTCAGAAGTTTACAGATATGTTAGCCAGGCCAACACCGT 360
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
QY 452 CAGACTTTCATCACTCAGTCATCAATTCCTGGGCCAGTATGAGTTTGGAGGCTGGAC 511
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
QY 361 CAGACTTTCATCACTCAGTCATCAATTCCTGGGCCAGTATGAGTTTGGAGGCTGGAC 420
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
QY 512 TTTGACTGGGAGTACCTGCTCTGCTGGAGCCCTCTCTCAGGACAAAGCATCTCTTCACT 571
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
```


Db 601 GTGACAAAATGCCAGAACCTGGATTGTTGTCACACCTTATGGCCCTACGACCTCCATGCG 660
QY 752 TCTGGAGGGCTACACTGGAGAGACAGCCCTCTCAAAATACCCGACTGACACCGGC 811
Db 661 TCTTGGGAAGGTCAAGGGACATACAGCCCTCTCAAGAGGCAAGAAGAGTGGT 720
QY 812 AGCAAGCGCTACTCAATGGATTATGTGTAAGTCTGTAAGGACAAATGAGCACCA 871
Db 721 GCAGCAGCCAGCTCAACTGAGTGTGTGTCACACAGTGGCTGAGAGGGGACCCCT 780
QY 872 GCTGAGAAGCTCATCTTGGATTCCCTACCTATGGACACAACTTCACTCTGAGCAACCC 931
Db 781 GCAGCAGAGCTGATCTTGGCATGCTACTACGAGCGCTCCCTTCACTACCTGCGCTCTCA 840
QY 932 TCCACACACTGGAATGGTGGCCCGCCACCTCTGGTGTGGTCTGCTGGGCGCCCTATGCCAAG 991
Db 841 TCAGACACACAGAGTGGGGCGCCAGCCACAGGCTCTGGACACTCCAGGCGCCCTTCACCAAG 900
QY 992 GAGTCTGGGATCTGGGCTTACTACGAGATCTGTACCTTCTGAAAAATGGAGCCACTCAG 1051
Db 901 GAAGGAGGATGCTGGCCCTACTATGAAGTCTGCTCT-----GGAAGGGGGGCCCAAA 954
QY 1052 GGATGGGATGCCCTCAGGAAGTGCCTTATGCTATCAGGGCAATGTGGGTGGCTAT 1111
Db 955 CAGAGAATCCAGGATCAGAAGGTGCCCTACATCTCCGGGACACACAGTGGGTGGCTT 1014
QY 1112 GACAACATCAAGAGCTTCATATTAAGGCTCAATGGCTTAAGCACAAACAAATTTGAGGC 1171
Db 1015 GATGATGTGGAGAGCTTCAAAACCAAGGTGAGCTATCTGAAGCAGAAGGAGTGGCGGG 1074
QY 1172 GCATGCTTGGGCCATGATCTGGATGACTTCACTGGCACTTCTGCAACACAGGCAAG 1231
Db 1075 GCGATGCTTGGCACTGACTTAGATGACTTTCGCGGCTTCTCCTGCAACACAGGCGCA 1134
QY 1232 TTTCCCTTAATCTCCACCTGAAGAGGCGCTCGGCTCGAGAGTGCAGATTCAGAGGCT 1291
Db 1135 TAGCCCTCATCAGACGCTACGGCAGGAAGTGAATGGGTAAAGCCTTAACCTGCTGTCA 1194
QY 1292 CCAGCTCAGCCCATG 1307
Db 1195 CATGTGAGGCCAGGTG 1210

RESULT 13
US-08-694-915-1
; Sequence 1, Application US/08694915
; Patent No. 5811535
; GENERAL INFORMATION:
; APPLICANT: Adamou, Julie
; APPLICANT: Kirkpatrick, Robert
; APPLICANT: Rosenberg, Martin
; TITLE OF INVENTION: HUMAN CARTILAGE GP39-LIKE GENE
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-2799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/694,915
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:

ATTORNEY/AGENT INFORMATION:
; NAME: Han, William T
; REGISTRATION NUMBER: 34,344
; REFERENCE/DOCKET NUMBER: ATG50017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5219
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1433 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; US-08-694-915-1

Query Match 22.7%; Score 369; DB 1; Length 1433;
Best Local Similarity 59.8%; Pred. No. 1.4e-98;
Matches 703; Conservative 0; Mismatches 445; Indels 27; Gaps 4;

QY 99 CAACCATGACAAAGCTTATTCTCTCACAGGCTCTTCTCTTATCTGTAATTTTCACGCTCG 158
Db 47 CCACCATGAGCAGAAAGTCTCTCTGGCAGGTGTAGTGTCTTCTGCTTCTCCAGGGAG 106
QY 159 GCTCTGCCCTACCAGCTGACATGCTACTTCCACCAACTGGGCCACAGTACCGGCCAGGCTGG 218
Db 107 GATCTGCCCTACAACTGGTTTGTCTACTTTACCAACTGGTCCAGGAGCCGCGAGAACCCAG 166
QY 219 GGGGCTTCATGCTGACACATCGACCCCTGCTGTGTACCCACCTGATCTAGCCCTTTG 278
Db 167 GAAATTTACCCCTGAGAAATATGTACCCCTTCTATGCTCTCATCTATTCATTCATTCG 226
QY 279 GTGGGAGGAGACAAAGAGATCACCACCATCAATGGAACGATGTGACTCTCTACCAAG 338
Db 227 CCAGCATCGAACAACAAGGTATCATCAAGACAAAGAGTGAAGTGTGCTCTACCA 286
QY 339 CTTTCAATGGCTGAAAAATAAGAACAGCAGCTGAAACTCTCTCTGGCCATTTGAGGCT 398
Db 287 CCATCAACAGTCTCAAAACCAAGAAATCCCAAAATTTCTCTGTCCATTTGAGGCT 346
QY 399 GGAATTTGGGACTGCCCTTTTCTACCTGCTGCTGAGAACCCGACACTT 458
Db 347 ACCTGTTGGTTCCAAAGGGTTCCACCCCTATGCTGATTTCTTACATCAGCTTGAAT 406
QY 459 TCATCACCTCAGTCATCAAAATCTCTGGCCAGTATGAGTTTGTACGGCTGGACTTTGACT 518
Db 407 TCATTAACCTCCATATCTCTGTTCTGAGGACCACTAATCTTGTGAGCTGGATGAAGCT 466
QY 519 GGGAGTACCCTGCTCTGCTGGAGGCGCTCTCAGGACAAAGCATCTCTTACCTGCTCTGG 578
Db 467 GGATCTACCCAGA-----TCAGAAAGAAACAACTCATTTTCTGCTGTGCTGA 511
QY 579 TGCAGGAATGCGTGAAGCTTTTGCAGGAGGCGCCAGCAGATCAACAAGCCGAGCTGA 638
Db 512 TTCATGAGTTAGCAGAAAGCTTTTCAGAAAGGACTTCAAAAATCCCAAGAAAGGCTTC 571
QY 639 TGGTCACTGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 698
Db 572 TCTTGTACTGGGGCGTATCTGAGGAGGCGCAATGATGTATACAGCTATCAAGTTGAGA 631
QY 699 AACTGTACAGTACCTGGACTACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 758
Db 632 AACTGGCAAAAGATCTGGATTTTCATCAACCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 691
QY 759 A-----GGGCTACACTGGAGAGAACAGCCCTCTTACAAATACCCGACTGACACCGCA 812
Db 692 AAAAGCCCTTATCTACTGGCCACACAGCCCTCTGAGCAAGGGGTGCGAGGAGGAGGCG 751

1052 GGATGGATGCCCTCAGGAAGTGCCTTATGCTATCAGGGCAATGTGTGGTTGGCTAT 1111
1078 ACCGGCTCCAGATCAGAGGTTCCCTACGCGAGTCAAGGGGAACAGTGGGTGGCTAT 1137
1112 GACAACATCAAGAGCTTCGATATTAAGGCTCAATGGCTTAAGCAACAATAATTTGGAGGC 1171
1138 GATGATGTGAAGAGTATGAGACCAAGGTTTCAGTTCTTAAAGAATTTAAACCTGGGAGGA 1197
1172 GCCATGGTGGGCCATTGATCTGGATGACTTCAGTGGCACTTTCTGCAACCAAGGCAAG 1231
1198 GCCATGATCTGGTCTTATGACATGGATGACTTCACCTGGCAAAATCCTGCAACCAAGGCCCT 1257
1232 TTTCCCTCAATCTCCACCTCGAAGAAGGCCCTCGGC 1267
1258 TACCCTCTGTCAAGCAGTCAAGAGAGGCTTGGC 1293

RESULT 15
PCT-US94-07754-4
; Sequence 4, Application PC/TUS9407754
; GENERAL INFORMATION:
; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
; TITLE OF INVENTION: ASSAY FOR YKL-40 AS A MARKER FOR
; TITLE OF INVENTION: DEGRADATION OF MAMMALIAN CONNECTIVE TISSUE MATRICES
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SPENSLEY HORN JUBAS & LUBITZ
; STREET: 1880 CENTURY PARK EAST, FIFTH FLOOR
; CITY: LOS ANGELES
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/07754
; FILING DATE: 08-JUL-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: HOWELLS, STACY L.
; REGISTRATION NUMBER: 34,842
; REFERENCE/DOCKET NUMBER: FD 3665
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/455-5100
; TELEFAX: 619/455-5110
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1681 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: YKL-40
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 135..1681
PCT-US94-07754-4

Query Match 20.9%; Score 339.8; DB 5; Length 1681;
Best Local Similarity 58.9%; Pred. No. 6e-90;
Matches 669; Conservative 0; Mismatches 442; Indels 24; Gaps 4;

125 ACAGGCTTGTCTTACTGATTTTCAGCTCGGCTCGCTCCCTACCAGCTGACATGCTAC 184
93 ACAGGCTTGTGTCTCTGTGCTGCTCCAGTGTGCTCTGCTGCATACAACTGGTCTGCTAC 152
185 TTCACCACTGGGCCAGTACCGGCGCGCTGGCGGCTTTCATGCTGCATACAACTGAC 244

153 TACACCAGCTGGTCCCGATACCGGGAAGCGATGGAGCTGCTTCCAGATGCCCTTGAC 212
245 CCCTGCTCTGTACCCACCTGATCTACGCTTTCCTGGGAGGAGGAGACACGAGATCACC 304
213 CGTTCCTGTGTACCCACATCATCTACAGCTTTGCCAATATAAGAACACGATACATCGAC 272
305 ACCATCGAATGGAACATGTGACTCTCTACCAAGCTTTTCAATGGCTTGAATAAAGAAC 364
273 ACCTGGAGTGAATGATGTGAGCTCTAGCGCATGCTCAACACATCAACACACGAAC 332
365 AGCAGCTGAAACTCTCTCTGGCCATTTGGAGGTGGAACCTTCGGGAGTCCCTTTCAC 424
333 CCCAAGCTGAAGACTCTCTGTCTGTGGAGGATGGAACCTTTGGGTCTTAAAGATTTTC 392
425 GCCATGTTTCTACTCTGAGAACCGCCACACTTTTCATCACTCAGTCATCAATTCCTG 484
393 AGATAGCTTCCCAACACCCAGAGTCCGCGGACTTTTCATCAAGTCACTGACCGCCATTC 452
485 CGCCAGTATGAGTTTGACGGGCTGGACTTTGACTGGGAGTACCTCTGCTCTCGTGGGAGC 544
453 CGCACCCTGCTTGTGATGGGCTGACCTTGCCTGCTCTACCTTGACGAGAGA----- 506
545 CTCTCAGGACAAAGCATCTCTTCACTGTCTCTGGTGCAGGAATGCGTGAAGCTTTGAG 604
507 -----GACAAACACCATTTTACACCTTATCAAGGAAATGAAGGCCGAATTTATA 557
605 CAGGAGGCCAAGCAGATCAACAAGCCAGGCTGATGCTCACTCTGTCAGTAGCTGCTGGC 664
558 AAGGAAGCCAGCCAGGGGAAAAGC---AGCTCTGCTCAGCGAGCCTGTCTGCGGGG 614
665 ATCTCCAATATCCAGTCTGGCTATGAGATCCCCAACTGTACAGTACCTTGGATACATC 724
615 AAGTCAACCATTCAGCAGCAGTATGACATTTGCCAAGATATCCCAACACCTTGGATTCAT 674
725 CATGTCATGACCTACGACCTCCATGCTCTCTGGAGGGCTACACTGAGAGAACAGCCCC 784
675 AGCATCATGACCTACGATTTTTCATGGCGCTGCGTGGGACCAACATCAGAGTCC 734
785 CTCTACAATATCCGACTGACACCGGAGCAACCGCTTACCTCAATGTTGATTTATGTCATG 844
735 CTGTTCCGAGGTGAGGAGTGAAGTCTCTGACAGATTGACAACTGACTATGCTGTG 794
845 AACTACTGGAAGGACAAATGGAGCACCAGCTGAGAAGCTCATCTGTTGATTTCCCTAC 904
795 GGCTACATGTTGAGGCTGGGGCTCTCTGCGAGTAAGCTGTGATGGGATCCCAACCTTC 854
905 GGACACAATTCATCTGAGCAACCCCTCCCAACACTGGAATTTGGTCCCGCCACCTCTGGT 964
855 GGGAGGAGCTTCACTCTG---GCTTCTCTGAGACTGGTGTTCAGCGCCCAATCTCAGGA 911
965 GCTGGTCTGCTGGGCCCTATGCAAGGAGTCTGGGATCTGGGCTTACTACGAGATCTGT 1024
912 CCGGGAATTCAGGCGGTTTCAACAGGAGGAGGACCTTCCCTACTATGATGATCTGT 971
1025 ACCTTCTCTGAAATAATGGAGCCACTCAGGGATGGGATGCCCTTACTAGAGTCTTATGCC 1084
972 GACTTCT---CCGCGGAGCCACAGTCCATAGAACCTCGCGCCAGCAGGTCCTCATGACC 1028
1085 TATCAGGGCAATGTGTGGGTTGGCTATGACAACATCAAGAGTTCGATATTAAGGCTCAA 1144
1029 ACCAAGGGCAACCAAGTGGTAGGATACGACGACGAGGAAAGCGCTCAAAAGAGGTGAG 1088
1145 TGCTTAAGCACACAATAATTTGGAGGCGCATGCTGGGCGCATGCTGAGTGGATGCTTC 1204
1089 TACTTAAGGATAGGACAGCTGGGAGGCGCCATGATGGGCGCTTGGACCTTGGATGACTTC 1148
1205 ACTGGCACTTTTCGAACCAAGGCAAGTTTCCCTTATCTCCACCTTGAAGAAGG 1259
1149 CAGGCTCTCTCTCGCGGCAAGATCTCGGCTTCCCTCTCACCATGCCATCAAGG 1203

Db 61 AGAATATAAAGCTCTGCGGACTGGTGTGCTGACTGCAACCAATGACAAAGCTTATCT 120
QY 121 CCTCAGAGTCTTGCTTATACGAATTTGACGTGCGCTCTGCTACACAGCTGACATG 180
Db 121 CCTCAGAGTCTTGCTTATACGAATTTGACGTGCGCTCTGCTACACAGCTGACATG 180
QY 181 CTACTTACCAACATGCGGCCAGTACCGGCCAGGCTGGGGCGCTTACGCTGCAACAT 240
Db 181 CTACTTACCAACATGCGGCCAGTACCGGCCAGGCTGGGGCGCTTACGCTGCAACAT 240
QY 241 CGACCCCTGCTCTGTACCCACCTGATCTACGCTTGTCTGGAGGAGACACAGAGAT 300
Db 241 CGACCCCTGCTCTGTACCCACCTGATCTACGCTTGTCTGGAGGAGACACAGAGAT 300
QY 301 CACCACCATCAATGGAACGATGTGACTCTCTACCAAGCTTTCAATGSCCTGAAAATAA 360
Db 301 CACCACCATCAATGGAACGATGTGACTCTCTACCAAGCTTTCAATGSCCTGAAAATAA 360
QY 361 GAACAGCAGCTGAAACTCTCTGAGCAATTTGAGGCTGGAACCTCGGCCACTGCCCTTT 420
Db 361 GAACAGCAGCTGAAACTCTCTGAGCAATTTGAGGCTGGAACCTCGGCCACTGCCCTTT 420
QY 421 CACTGCCATGTATGAGTGTGAGCGGCTGGACTTTGACTGGAGTACCTCAGTCATCAAT 480
Db 421 CACTGCCATGTATGAGTGTGAGCGGCTGGACTTTGACTGGAGTACCTCAGTCATCAAT 480
QY 481 CCTGGCCAGTATGAGTGTGAGCGGCTGGACTTTGACTGGAGTACCTCAGTCATCAAT 540
Db 481 CCTGGCCAGTATGAGTGTGAGCGGCTGGACTTTGACTGGAGTACCTCAGTCATCAAT 540
QY 541 GAGCCCTCTCAGACCAAGCATCTCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
Db 541 GAGCCCTCTCAGACCAAGCATCTCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
QY 601 TGACGAGGAGCCAGCAGATCAACAGCCAGCTGATGCTCAGTCTGCTGCTGCTGCTGCTGCT 660
Db 601 TGACGAGGAGCCAGCAGATCAACAGCCAGCTGATGCTCAGTCTGCTGCTGCTGCTGCTGCT 660
QY 661 TGGCATCTCCAAATATCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Db 661 TGGCATCTCCAAATATCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
QY 721 CATCCATCTCATGACCTACGACCTTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db 721 CATCCATCTCATGACCTACGACCTTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
QY 781 CCCCTCTACAAATACCGACTGACCGGAGCAGCAACGCTTACCTCAATGCTGCTGCTGCTGCT 840
Db 781 CCCCTCTACAAATACCGACTGACCGGAGCAGCAACGCTTACCTCAATGCTGCTGCTGCTGCT 840
QY 841 CATGAATCTGGAAGGACAAATGAGCAGCAGCTGAGAGCTCATGCTGCTGCTGCTGCTGCTGCT 900
Db 841 CATGAATCTGGAAGGACAAATGAGCAGCAGCTGAGAGCTCATGCTGCTGCTGCTGCTGCTGCT 900
QY 901 CTATGGACACAACTTCTGCTGACCAACCCCTGCAACACTGGAATTTGGTCCGCCCTCCTC 960
Db 901 CTATGGACACAACTTCTGCTGACCAACCCCTGCAACACTGGAATTTGGTCCGCCCTCCTC 960
QY 961 TGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
Db 961 TGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
QY 1021 CTGTACCTCTCTGAAAATGAGCCACTCAGGATGGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
Db 1021 CTGTACCTCTCTGAAAATGAGCCACTCAGGATGGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
QY 1081 TGGCTATCAGGCAATGTGTGGTGGCTATGCAACATCAAGAGCTTCGATATTAAGG 1140
Db 1081 TGGCTATCAGGCAATGTGTGGTGGCTATGCAACATCAAGAGCTTCGATATTAAGG 1140
QY 1141 TCAATGGCTTAAACACAAATTTGGAGGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
Db 1141 TCAATGGCTTAAACACAAATTTGGAGGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200

QY 1201 CTTCACTGGCACTTTCTGCAACAGGCAAGCTTTCCCTAATCTCCACCTGAAGAAGGC 1260
Db 1201 CTTCACTGGCACTTTCTGCAACAGGCAAGCTTTCCCTAATCTCCACCTGAAGAAGGC 1260
QY 1261 CCTCGGCTCGAGAGTGCAGTGTGACGGCTCCAGCTCAGCCCATTTAGCCCAATAACTGC 1320
Db 1261 CCTCGGCTCGAGAGTGCAGTGTGACGGCTCCAGCTCAGCCCATTTAGCCCAATAACTGC 1320
QY 1321 TGGTCCCACTGGCAGCGGGAACGGAGGGGAGTAGCAGCTCTGAGGAGCTCGGAGG 1380
Db 1321 TGGTCCCACTGGCAGCGGGAACGGAGGGGAGTAGCAGCTCTGAGGAGCTCGGAGG 1380
QY 1381 CAGTGGATTTCTGCTGTGACAGCAACGGGCTCTACCCCTGGCAATAACAGAAATGC 1440
Db 1381 CAGTGGATTTCTGCTGTGACAGCAACGGGCTCTACCCCTGGCAATAACAGAAATGC 1440
QY 1441 CTTCTGCACTGCTGATGAGTGTGACGACGAACTGACGAGGCGGCTGCTGCTGCTGCT 1500
Db 1441 CTTCTGCACTGCTGATGAGTGTGACGACGAACTGACGAGGCGGCTGCTGCTGCTGCT 1500
QY 1501 CGACACCACTGCTGATGAGTGTGACGACGAACTGACGAGGCGGCTGCTGCTGCTGCTGCT 1560
Db 1501 CGACACCACTGCTGATGAGTGTGACGACGAACTGACGAGGCGGCTGCTGCTGCTGCTGCT 1560
QY 1561 GTTCCAGTCTCTTTTCTGCTGAGGACATGTTGCCCTTACCTAAAGTCTGCAATAAAATCAG 1620
Db 1561 GTTCCAGTCTCTTTTCTGCTGAGGACATGTTGCCCTTACCTAAAGTCTGCAATAAAATCAG 1620
QY 1621 CAGTC 1625
Db 1621 CAGTC 1625

RESULT 2

US-10-004-219B-3
; Sequence 3, Application US/10004219B
; Publication No. US20030087414A1
; GENERAL INFORMATION:
; APPLICANT: Macrozyme
; APPLICANT: Aerts, Johannes M.F.G.
; APPLICANT: Boot, Rolf G.
; TITLE OF INVENTION: A mammalian mucinase, its recombinant production, and its use in therapy or prophylaxis against diseases in which mucus is involved or infection diseases
; TITLE OF INVENTION: its use in therapy or prophylaxis against diseases in which mucus is involved or infection diseases
; FILE REFERENCE: 2183-5136US
; CURRENT APPLICATION NUMBER: US/10/004,219B
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1525
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Combined DNA/RNA Molecule: mouse
; OTHER INFORMATION: AMCase cDNA sequence and deduced amino acid
; OTHER INFORMATION: sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: mouse AMCase
; OTHER INFORMATION: cDNA sequence and deduced amino acid sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1419)
US-10-004-219B-3

Query Match 64.1%; Score 1041; DB 9; Length 1525;
Best local similarity 83.4%; Pred. No. 0;
Matches 1198; Conservative 0; Mismatches 230; Indels 9; Gaps 1;
QY 104 ATCAACAAGCTTATCTCTCAGGCTGCTGCTTATCTATCTGCTGCTGCTGCTGCTGCTGCTGCT 163
Db 1 ATGCCCAAGCTTATCTCTCAGGCTGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 60

Db 1027 GACTTCTCTCCGCGAGCCACAGTCCATAGAACCTCGGCCAGCAGGTCCCTATGCC 1083
QY 1085 TATCAGGGAATGTGGGTGGCTATCAACATCAAGAGCTTCGATATTAAGGCTCAA 1144
Db 1084 ACCAAGGGAACAGTGGGTAGATACGACGACCAAGAGCGTCAAAAGCAAGGTGCAG 1143
QY 1145 TGGCTTTAAGCACAAACAAATTTGGAGCGGCATGGTCTGGGCGCAATGATCTGGATGACTTC 1204
Db 1144 TACCTAAGGATAGGAGCTGGCAGCGGCATGGTATGGCCCTGGACCTGGATGACTTC 1203
QY 1205 ACTGGCAGCTTTCTGCAACAGGCAAGTTTCCCTTAATCTCCACCCTGAAGAAGG 1259
Db 1204 CAGGGCTCTTCTGCGCGCAGGATCTGCGTCTCCCTCTCACCAATGCCATCAAGG 1258

RESULT 5

US-09-822-830A-402
; Sequence 402, Application US/09822830A
; Patent No. US20020142952A1
; GENERAL INFORMATION:
; APPLICANT: Genetics Institute, Inc.
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fechtel, Kim
; APPLICANT: Agostino, Michael J.
; APPLICANT: Howes, Steven H.
; APPLICANT: Resnick, Richard J.
; APPLICANT: Gulukota, Kamalakari
; APPLICANT: Graham, James R.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
; FILE REFERENCE: GIN 6402
; CURRENT APPLICATION NUMBER: US/09/822,830A
; CURRENT FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/195,604
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 631
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 402
; LENGTH: 1391
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1353
; OTHER INFORMATION: n=a,c,g, or t
US-09-822-830A-402

Query Match 21.4%; Score 348.4; DB 10; Length 1391;
Best Local Similarity 59.2%; Pred. No. 6.6e-105;
Matches 680; Conservative 0; Mismatches 441; Indels 27; Gaps 4;
QY 126 CAGGCTCTGCTTATACATGAATTTGCGAGCTCGGCTCTGCTACAGCTGACATGCTACT 185
Db 35 CAGGTAGTGGTCTTCTGCTGCTTCTCCAGGAGGATCTGCTTACAACTGGTTGGGACT 94
QY 186 TCACCAACTGGGCCCCAGTACCGCCAGGCTGGGGGCTTCATGCTGCAACATCGACC 245
Db 95 TTACCAACTGGTCCAGGACCGGCGAGCAACGAGAAATTCACCCCTGAGATATGACC 154
QY 246 CCGCTCTGTATCCCACTGATCTACGCCCTTTGCTGGGAGGCAACACAGAGATCACCA 305
Db 155 CTTCTCTATGCTCTCATCTCATCTATTTCATTCGCCAGCATCGAAACAAACAGGTATCA 214
QY 306 CCATCGAATGGACGATGTGACTCTCTACCAAGCTTTCAATGGCTGGAAATTAAGACA 365
Db 215 TCAAGGACAGAGTGAAGTGTGCTCTACAGACCATCAACAGCTCTCAAAACCAAGAAC 274
QY 366 GCAGCTGAAACTCTCTGGCCATGGAGGCTGGAACTTTGCGGACTGCCCTTTCACTG 425
Db 275 CCAACTGAAATTCCTTGTCTCCATTGGAGGGTACCTGTTGGTCCAAAGGGTTCACC 334
QY 426 CCATGTTTCTACTCTCTGAGAACCGCCAGACTTTTCATCAGCTCAGTCAAAATTCCTGC 485

Db 335 CTATGGTGGATTTCTTACATCAGCTTGGATTCATTAACCTCATTAATCCTGTTCTGA 394
QY 486 GCCAGTATGATGTTGAGGGCTGACTTTTACCTGGGAGTACCTGGCTCTCTCTGGGAGCC 545
Db 395 GGAACCATAACTTTGATGGACTGGATGAAGCTGGATCTACCCAG----- 439
QY 546 CTCTCAGGACAAAGCATCTCTTACCTCTCTGTCAGAGAAATCGCTGAAGCTTTTGAAGC 605
Db 440 ATCAGAAAGAAACACTCATTTCACTCTGCTGATTCATGATAGTGGAGAGAGCCCTTCAGA 499
QY 606 AGGAGGCCAAGCAGATCAACAGCCAGGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 665
Db 500 AGGACTTCAAAAATCCACCAAGGAAAGGCTTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTG 559
QY 666 TCTCCAAATCCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 725
Db 560 GGCAAATGATTTGATTAACAGCTATCAAGTTGAGAACTGGCAAAAGATCTGGATTTATCA 619
QY 726 ATGCTATGACTACGACTCCTCATGGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 779
Db 620 ACCTCTGCTCTTTGACTTCCATGGTCTTGGGAAAGCCCTTATCACTGGCCACAACA 679
QY 780 GCCCTCTTACAAATACCCGACTGACACCGGAGCAAGCCCTACCTCAATGTGGATATG 839
Db 680 GCCCTCTGAGCAAGGGGTGGCAGGACAGAGGGCCAAAGCTCTACTACAATGTGGAATATG 739
QY 840 TCATGAACCTGGAAGGCAATGGAGCACCAGCTGAGAGCTCATCTGTTGGATTTCCCTA 899
Db 740 CTGTGGGTACTGTATACATTAAGGAATGCCATCAGAGAAGGTGGTCTATGGCATCCCA 799
QY 900 CCTATGACACAACTTCTATCTGAGCAACCCCTCCAAACACTGGAATTTGGTCCCTACCT 959
Db 800 CATATGGGCACTCTTCACTGCTCTGCAAGAACCCG---TGCGGCCCCCTGCTGCT 856
QY 960 CTGGTGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1019
Db 857 CTGGCTCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 916
QY 1020 TCTGTACTTCTCTGAAATTTGGAGCCACTCAGGATGGGATGCTGCTGCTGCTGCTGCTGCT 1079
Db 917 TCTGCCAGTTTCTGAAA---GGAGCCCAAGATCAGAGGCTCCAGAGTTCAGAGTTCCCT 973
QY 1080 ATGCTATCAGGCAATGTGTGGTGTGGCTATGACAAACATCAAGAGCTTCGATATTAAGG 1139
Db 974 ACCAGTCAAGGGGACCCGCTGGTGGCTTATGATGATGATGATGATGATGATGATGATGAT 1033
QY 1140 CTCAATGGCTTAAAGCACAACAAATTTGGAGGCGCCATGGTCTGCGCCATTTGATCTGATG 1199
Db 1034 TTCAGTTCTTAAAGAAATTTAAAGCTGGGGGTGCTTGTGCTGCTGCTGCTGCTGCTGCTG 1093
QY 1200 ACTTCACTGCACTTCTTCAACAGGCAAGTTTCCCTTAATCTCCACCTGGAAGAGG 1259
Db 1094 ACTTCACTGCAAAATCTGCAACGGGGCCCTTCCCTCTTGTGCTGCTGCTGCTGCTGCTG 1153
QY 1260 CCCTCGGC 1267
Db 1154 GCCTTGGC 1161

RESULT 6

US-09-960-352-678
; Sequence 678, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352

; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 678
; LENGTH: 449
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 03-LIB34-081-Q1-E1-A11
US-09-960-352-678

Query Match 21.0%; Score 342; DB 10; Length 449;
Best Local Similarity 87.2%; Pred. No. 4.6e-103;
Matches 375; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 139 TATACGTAATTCAGCTCGGCTCTGCTACAGCTGACATGCTACTTCCACCACTGGGC 198
DB 139 TATACGTAATTCAGCTCGGCTCTGCTACAGCTGACATGCTACTTCCACCACTGGGC 198
QY 199 CCAGTACGGCCAGGCTGGGGCTTCATGCTGACAACTGACACCCCTGCCCTCTGPAC 258
DB 199 CCAGTACGGCCAGGCTGGGGCTTCATGCTGACAACTGACACCCCTGCCCTCTGPAC 258
QY 259 CCACCTGATCTACGCTTTCGCTGGGAGGCGAAGCAACAGATCACCACCTGATGGAA 318
DB 259 CCACCTGATCTACGCTTTCGCTGGGAGGCGAAGCAACAGATCACCACCTGATGGAA 318
QY 319 CGATGCTACTCTACCAAGCTTCAATGGCTGAAAAATAAGAACAGCCAGCTGAAAAAC 378
DB 319 CGATGCTACTCTACCAAGCTTCAATGGCTGAAAAATAAGAACAGCCAGCTGAAAAAC 378
QY 200 TGATGTGCTCTCTATAGTCTTCTCAATGACTGAAAAAAGAACAGCCAGCTGAAAAAT 259
DB 200 TGATGTGCTCTCTATAGTCTTCTCAATGACTGAAAAAAGAACAGCCAGCTGAAAAAT 259
QY 379 TCTCTGCCCATTTGAGGCTGGAATCTGGGAGCTGCCCCCTTTCATGCTGCTGTTCTPAC 438
DB 379 TCTCTGCCCATTTGAGGCTGGAATCTGGGAGCTGCCCCCTTTCATGCTGCTGTTCTPAC 438
QY 439 TCCGTGAGAACCGCAGACTTTCATCACCTCAGTCAATCAATTCCTGCCCGCAGTATGATT 498
DB 439 TCCGTGAGAACCGCAGACTTTCATCACCTCAGTCAATCAATTCCTGCCCGCAGTATGATT 498
QY 499 TGACGGGCTGGACTTTGACTGGGAGTACCTTGGCTCTGCTGGGAGCCCTCCTCAGGACAA 558
DB 499 TGACGGGCTGGACTTTGACTGGGAGTACCTTGGCTCTGCTGGGAGCCCTCCTCAGGACAA 558
QY 559 GCATCTCTTC 568
DB 559 GCATCTCTTC 568

RESULT 7

US-09-215-077A-4
; Sequence 4, Application US/09215077A
; Patent No. US20020031793A1

; GENERAL INFORMATION:
; APPLICANT: PRICE, PAUL A.

; APPLICANT: JOHANSEN, JULIA S.
; TITLE OF INVENTION: ASSAY FOR YKL-40 AS A MARKER FOR DEGRADATION OF

; FILE REFERENCE: 407T-895411US
; CURRENT APPLICATION NUMBER: US/09/215,077A

; CURRENT FILING DATE: 1998-12-18
; PRIOR APPLICATION NUMBER: 08/581,527

; PRIOR FILING DATE: 1996-04-17
; PRIOR APPLICATION NUMBER: 08/089,989

; PRIOR FILING DATE: 1993-07-09
; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4

; LENGTH: 1681
; TYPE: DNA

; ORGANISM: Artificial Sequence
; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: cDNA nucleotide
; OTHER INFORMATION: sequence for the coding region of the gene for

; OTHER INFORMATION: YKL-40.
US-09-215-077A-4

Query Match 20.8%; Score 338.2; DB 10; Length 1681;
Best Local Similarity 58.9%; Pred. No. 1.8e-101;
Matches 668; Conservative 0; Mismatches 443; Indels 24; Gaps 4;
QY 125 ACAGGCTCTGCTTATCTACTGAATTTGACAGCTCGGCTCTGCTACAGCTGACATGCTPAC 184
DB 125 ACAGGCTCTGCTTATCTACTGAATTTGACAGCTCGGCTCTGCTACAGCTGACATGCTPAC 184
QY 185 TTCACCACTGGGCCAGTACCGGCCAGGCTTGGGGCTTTCATGCTCCCTGACACATCAGC 244
DB 185 TTCACCACTGGGCCAGTACCGGCCAGGCTTGGGGCTTTCATGCTCCCTGACACATCAGC 244
QY 153 TACACCACTGGTCCAGTACCGGGAAGGCGATGGGAGCTGCTTCCCAAGATGCCCTTAC 212
DB 153 TACACCACTGGTCCAGTACCGGGAAGGCGATGGGAGCTGCTTCCCAAGATGCCCTTAC 212
QY 245 CCCTGCTCTCTACCCACCTCTATCTACGCTTGTGGAGGCGAAGCAACAGATGATCACC 304
DB 245 CCCTGCTCTCTACCCACCTCTATCTACGCTTGTGGAGGCGAAGCAACAGATGATCACC 304
QY 213 CGCTTCTGCTACCCACATCTATACAGCTTGTCCCAATATAGCAAGATGATCAGTAC 272
DB 213 CGCTTCTGCTACCCACATCTATACAGCTTGTCCCAATATAGCAAGATGATCAGTAC 272
QY 305 ACCATCGAATGGAACGATGTGACTCTCTACCAAGCTTTCATGCTGCTGCTGCTGCTG 364
DB 305 ACCATCGAATGGAACGATGTGACTCTCTACCAAGCTTTCATGCTGCTGCTGCTGCTG 364
QY 273 ACCTGGGAGTGAATGATGTGACGCTCTACGGCATGCTCAACACACTCAACACAGCAAC 332
DB 273 ACCTGGGAGTGAATGATGTGACGCTCTACGGCATGCTCAACACACTCAACACAGCAAC 332
QY 365 AGCCAGCTGAAACCTCTCTGCGCATTTGGAGGCTTGGAACTTCTGCTGCTGCTGCTG 424
DB 365 AGCCAGCTGAAACCTCTCTGCGCATTTGGAGGCTTGGAACTTCTGCTGCTGCTGCTG 424
QY 333 CCCAACCTTGAAGACTCTTCTGCTGCGAGGATGGAATTTGGGCTCTCAAGATTTTTC 392
DB 333 CCCAACCTTGAAGACTCTTCTGCTGCGAGGATGGAATTTGGGCTCTCAAGATTTTTC 392
QY 425 GCCATGTTTCTACTCTCTGAGAACCGCCAGACTTTTCATCAGCTCAGTCAATCAATTC 484
DB 425 GCCATGTTTCTACTCTCTGAGAACCGCCAGACTTTTCATCAGCTCAGTCAATCAATTC 484
QY 393 AAGATAGCTCCACACCCAGAGTGGCGGACTTTCATCAAGTCAAGTACCCGCTTCTG 452
DB 393 AAGATAGCTCCACACCCAGAGTGGCGGACTTTCATCAAGTCAAGTACCCGCTTCTG 452
QY 485 CGCAGATGATGTTGAGGCTGAGCTTTGACTGGGAGTACCTTGGCTCTGCTGCTGCTG 544
DB 485 CGCAGATGATGTTGAGGCTGAGCTTTGACTGGGAGTACCTTGGCTCTGCTGCTGCTG 544
QY 453 CGCACCTTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 506
DB 453 CGCACCTTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 506
QY 545 CCTCTCAGGACAGCATCTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 604
DB 545 CCTCTCAGGACAGCATCTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 604
QY 507 -----GACAAACACCATTTTACCACCTTATCAAGGAATGAAAGGCGCAATTTATA 557
DB 507 -----GACAAACACCATTTTACCACCTTATCAAGGAATGAAAGGCGCAATTTATA 557
QY 605 CAGGAGGCAACAGATCAACAAAGCCAGGCTGATGCTGCTGCTGCTGCTGCTGCTGCTG 664
DB 605 CAGGAGGCAACAGATCAACAAAGCCAGGCTGATGCTGCTGCTGCTGCTGCTGCTGCTG 664
QY 558 AAGGAGCCAGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 614
DB 558 AAGGAGCCAGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 614
QY 665 ATCTCAATATCCAGTCTGCTGATGAGATCCCCCACTGCTCAGTACAGTACATC 724
DB 665 ATCTCAATATCCAGTCTGCTGATGAGATCCCCCACTGCTCAGTACAGTACATC 724
QY 615 AAGGTCACTTACAGCAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 674
DB 615 AAGGTCACTTACAGCAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 674
QY 725 CATGCTAGCTACGACCTCCATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 784
DB 725 CATGCTAGCTACGACCTCCATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 784
QY 675 AGCATCATGACCTACGATTTTTCATGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 734
DB 675 AGCATCATGACCTACGATTTTTCATGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 734
QY 785 CTCTACAAATACCCGACTGACACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 844
DB 785 CTCTACAAATACCCGACTGACACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 844
QY 735 CTCAGGCGAGTCTAGGAGGATGCAAGTCTCTGAGATTCAGACACTGATGCTGCTG 794
DB 735 CTCAGGCGAGTCTAGGAGGATGCAAGTCTCTGAGATTCAGACACTGATGCTGCTG 794
QY 845 AACTACTGGAAGCAATGGAGCACCAGCTGAGAGCTCACTGCTGCTGCTGCTGCTGCTG 904
DB 845 AACTACTGGAAGCAATGGAGCACCAGCTGAGAGCTCACTGCTGCTGCTGCTGCTGCTG 904
QY 795 GGGTACATGTTGAGGCTGGGGCTCTCCAGTAACTGCTGCTGCTGCTGCTGCTGCTG 854
DB 795 GGGTACATGTTGAGGCTGGGGCTCTCCAGTAACTGCTGCTGCTGCTGCTGCTGCTG 854
QY 905 GGACACACTTCACTCTGAGCAACCCCTCCAACTGGAATTTGCTGCTGCTGCTGCTGCT 964
DB 905 GGACACACTTCACTCTGAGCAACCCCTCCAACTGGAATTTGCTGCTGCTGCTGCTGCT 964
QY 855 GGGAGGAGCTTCACTCTG---GCTTCTTCTGAGACTGCTGCTGCTGCTGCTGCTG 911
DB 855 GGGAGGAGCTTCACTCTG---GCTTCTTCTGAGACTGCTGCTGCTGCTGCTGCTG 911
QY 965 GCTGCTCTGCTGGGCTTATGCCAAGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1024
DB 965 GCTGCTCTGCTGGGCTTATGCCAAGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1024
QY 912 CCGGGAATTCAGGCCCGCTTCCACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 971
DB 912 CCGGGAATTCAGGCCCGCTTCCACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 971
QY 1025 ACCTTCTGAAAAATGGAGCCTCTCAGGAGTGGATGCTGCTGCTGCTGCTGCTGCTG 1084
DB 1025 ACCTTCTGAAAAATGGAGCCTCTCAGGAGTGGATGCTGCTGCTGCTGCTGCTGCTG 1084
QY 972 GACTTCTCT---CCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1028
DB 972 GACTTCTCT---CCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1028
QY 1085 TATCAGGGCAATGCTGGGTTGGCTATGACAACTCAAGAGCTTCGATATTAAGGCTCAA 1144
DB 1085 TATCAGGGCAATGCTGGGTTGGCTATGACAACTCAAGAGCTTCGATATTAAGGCTCAA 1144
QY 1029 ACCAAGGGCAACCAAGTGGGTAGGATACGACGACGAGGAGGAGGAGGAGGAGGAGG 1088
DB 1029 ACCAAGGGCAACCAAGTGGGTAGGATACGACGACGAGGAGGAGGAGGAGGAGGAGG 1088

Qy	1145	TGGCTTAAACACAAACAATTTGGAGGCGCCATGGTCTGGGCGCAATTGATCTGGGATGACTTC	1204
Db	1089	TACCTGAAGGATAGGCAGCTGGCAGGCGCCATGGTATGGGCCCCCTGGACCTTGGATGACTTC	1148
Qy	1205	ACTGGCACTTTCTGCACACAGGCGAAGTTTTTCCCCCTAATCTCCACCTTGAAGAAGG	1259
Db	1149	CAGGGTCTCTTCTGGCGCCAGGATCTGCGCTTCCCTCTACCAATGCCATCAAGG	1203

RESULT 8

```

US-09-262-213A-4
; Sequence 4, Application US/09262213A
; Patent No. US20020090638A1
; GENERAL INFORMATION:
; APPLICANT: PRICE, PAUL
; APPLICANT: JOHANSSEN, JULIA
; TITLE OF INVENTION: ASSAY FOR YKL-40 AS A MARKER FOR DEGRADATION OF
; TITLE OF INVENTION: MAMMALIAN CONNECTIVE TISSUE MATRICES
; FILE REFERENCE: 407T-895412US
; CURRENT APPLICATION NUMBER: US/09/262,213A
; CURRENT FILING DATE: 1999-03-04
; PRIOR APPLICATION NUMBER: US 08/089,989
; PRIOR FILING DATE: 1993-07-09
; PRIOR APPLICATION NUMBER: PCT/US94/07754
; PRIOR FILING DATE: 1994-07-08
; PRIOR APPLICATION NUMBER: US 08/581,527
; PRIOR FILING DATE: 1996-04-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO. 4
; LENGTH: 1681
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-262-213A-4

```

Query Match	20.8%;	Score 338.2;	DB 10;	Length 1681;
Best Local Similarity	58.9%;	Pred. No. 1.8e-101;		
Matches	668;	Conservative 0;	Mismatches 443;	Indels 24; Gaps 4;
Qy	125	ACAGGCTTCTGCTTATAC	TGAATTTCCAGCTCGGCT	TCGCCTACCAAGCTGACATGCTAC 184
Db	93	ACAGGCTTTGTGGTCCCTG	GTGCTGCCTCCAGTGCCTG	CTCGATACAACTGGTCTGCTAC 152
Qy	185	TTCAACCAACTGGGCCAGT	ACCGGCCAGGCTGGGGCCT	TCATGCTGACAAACATCGAC 244
Db	153	TACACCAAGCTGGTCCCA	GTACCGGGAAGCGATGGG	AGCTGCTCCACAGATGCCCTTGAC 212
Qy	245	CCCTGCTCTGTACCCACCT	GTATCTACGCCCTTTTCT	GGGAGGCAACAACGAGATCAC 304
Db	213	CGCTTCCTGTACCCACAT	CTACTACAGCTTTTGCCA	ATATAAGCAACGATCACATCGAC 272
Qy	305	ACCATCGAATGGGAAGAT	GTGACTCTCTACCAAGCT	TTTCAATGGCCTGAAAAATAGAAC 364
Db	273	ACCTGGGAGTGGATGATG	TGACGCTCTAGGGCATG	CTCAACACACTCAACAACACGAAC 332
Qy	365	AGCCAGCTGAAACTCTCT	CGGCATTTGGAGGCTTGG	AACCTCGGGAGCTGCCCTTTTCACT 424
Db	333	CCCAACCTGAAGACTCT	CTTGTCTGTCGGAGGAT	TGGAACCTTGGGCTCTCAAAGATTTTCC 392
Qy	425	GCCATGTTTCTACTCTG	AGAACCGCCAGAGCTTTC	ATCACCCTCAGTCATCAAAATTCCTG 484
Db	393	AGATAGCCTCCAAACACC	AGATGTCGGGGACTTTT	CATCAAGTCAGTACGACCGCCATTTCTG 452
Qy	485	CGCCAGTATGAGTTTGAC	GGGCTGGACTTTGACPG	GGAGTACCCCTGGCTCTCGTGGGAGC 544
Db	453	CGCACCATGGCTTTGAT	GGCGGTGACCTTGCCT	GGCTCTACCCCTGACAGCGAGA----- 506
Qy	545	CCTCCTCAGGACAGCAT	CTCTCTCCACTGTCTGT	GTCAGGAAATGCCGTGAAGCTTTTGAG 604
Db	507	-----GACAAACACCA	TTTACACCCCTAATCA	AGGAAATGAAGGCCGAATTTATA 557
Qy	605	CAGGAGGCCAAGCAGAT	CAACAAGCCAGGCTGAT	TGTTCACTGTCGAGTAGCTGCTGGC 664

DECLASS

```

RESULTS 9
US-09-765-231A-8
; Sequence 8, Application US/09765231A
; Patent No. US20020119452A1
; GENERAL INFORMATION:
; APPLICANT: Searle/Monsanto
; APPLICANT: Philippard, Deborah
; APPLICANT: Vasanthakamur, Geetha
; APPLICANT: Dotson, Stanton
; APPLICANT: Ma, Xiao-Jun
; TITLE OF INVENTION: Osteoarthritis tissue-derived nucleic acids, polypeptides,
; TITLE OF INVENTION: vectors, and cells
; FILE REFERENCE: SO-3221 PR
; CURRENT APPLICATION NUMBER: US/09/765,231A
; CURRENT FILING DATE: 2001-01-18
; NUMBER OF SEQ ID NOS: 82
; SEQ ID NO 8
; LENGTH: 1474
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-765-231A-8

```

	Query Match	17.3%	Score 281.2;	DB 10;	Length 1474;
	Best Local Similarity	59.4%;	Pred. No. 1.6e-85;		
	Matches 667;	Conservative 0;	Mismatches 423;	Indels 32;	Gaps 10;
Qy	158	GGCTCTGCCCTACAGCTGCATGCTACTTCCACCAACTGGG-CCCAGTACCGGCCGAGCGCT	216		
Db	128	GGATCTGCCCTACAACTGGTTTGGTACTTTACCAACTGGGTCCCGAGACCGGAGGAACCC	187		

Qy	217	GGGGCGCTTCATGCTTGAC - AACATCGAGCCCTCGCTGTGTACCCACCTGATCTACGGCT	275
Db	188	AGAAAAATTCACCCCTCGAGGAATATTGACCCCTTCCTATGCTCTCTCATCTCATATTTCAT	247
Qy	276	T - TCGTGGGAGGAGAACACGAGATCACACCAATCGAATGGAATGGAAGATGCACCTCTCTAC	334
Db	248	TGCGCCAGCATCGAAAACAACAGGTTATCATCAAGGACAAGATGAAGTGAATGCTCTCTAC	307
Qy	335	CAAGCTTTCAA - TGCCCTGAAAAATAAGAACAGCAGCTGAAAACTCTCTCGGGCATTTGG	393
Db	308	CAGACCATCAACAGCTCTCTAAAACCAAGAATCCCAAACCTGAAAATCTCTTGTCCATTGG	367
Qy	394	AGGCTGAACCTCGGGACTGCCCTTTTCATGCGCATGGTTTCTACTCTCTGAGAACCGCCA	453
Db	368	AGGTAACCTGTTTGGTTTCCAAAGGGTTCCACCCCTATGGTGTATCTTCTATCATCACGCTT	427
Qy	454	GACTTTTCATCACCTTCAGTCATCAAAATCTCGCGCAGTATGAGTTTGAAGGGCTGGACTT	513
Db	428	GGAAATTCATTAACTTCATATCTCTTCTGAGGAACCATAACTTTGATGGAGCTGGATGT	487
Qy	514	TGACTGGGAGTACCCTTGCTCTCTGTGGAGCCCTCCTCAGGACAAGCATCTCTTCACTGT	573
Db	488	AAGCTGATCTACCCAG - - - - - ATCAGAAAAGAAAACACTCATTTCACTGT	532
Qy	574	CCTGGTGCAGAAATCGTGAAGCTTTTGACGAGGAGCCCAAGCAGATCAACAAGCCAG	633
Db	533	GCTGATTCATGAGTTACGAGNAGCCTTTCAGAGAGCTTCACAAAATCCACCAAGGAAG	592
Qy	634	GCTGATGGTCACCTGCTGCA - GTAGCTGCTGGCATCTCCAATATCCAGTCTGGCTATGAGA	692
Db	593	GCTTCTCTTACTCGGGGGTATCTCGAGGAGGCAAAATGATTGATAACAGATATCAAG	652
Qy	693	TCCCCAACTGTCAAGTACTGGACTACATCCATGTATGATGACCTACGACCTCCATGGCT	752
Db	653	TTGAGAAACTGGCAAAAGATCTGGATTTCAACAACCTCCTGCTCTTGACTTTCATGGGT	712
Qy	753	CCTGGGA - - - - - GGGGTACACTGGAGAGAACAGCCC - CCTCTACAATAACCCGACTGAC	805
Db	713	CTTGGGAAAAGCCCTTTATCATTGCGCACAAACAGCCCTGCTGAGCAAGGGGTGCGAGAC	772
Qy	806	ACGGCAGCAACGCTACCTCAATGTGGATTATGTTCATGAATCTGGAAGGACAATGGA	865
Db	773	AGAGGCCCAAGCTCTACTACAATGTGGAAATATCTGTGGGGTATCTGGATACATAAGGGA	832
Qy	866	GCACCACTGAGAACTCATCTGTGGATTCCTTACCTATGAGACACAACCTCATCTCCTGAGC	925
Db	833	ATGCCATCAGAGAAGTGGTCATGGGCATCCCCACATATGG - - GGCATCTTCACACTG	890
Qy	926	AACCCCTCCAACACTGGAAATGGTGTGCCCCACCTCTGCTGTGCTGTCTGTGGCCCTAT	985
Db	891	GCCTCTGCAGAAACCAACCTGTGGGGGCCCTGCTCTGCGCCTGGAGCTGTGGACCCATC	950
Qy	986	GCCAAGGAGTCTGGGATCTGGCTTACTACGAGATCTGTACCTTCTCTGAAAAATGGAGCC	1045
Db	951	ACAGAGTCTTCAGGCTTCTCTGGCCCTATATGAGATCTGCGCAGTTTCTTGGAAA - - - GGAACC	1007
Qy	1046	ACTCAGGAGTGGATGCCCCCTCAGGAAGTGCCCTTATGCCATCATTAGGGCAATGTGTGGGTT	1105
Db	1008	AAGATCAGCGGCTCCAGGATCAGCAGGTTCCCTACCGAGTCAAGGGGACCACTGGGTG	1067
Qy	1106	GGCTATGACAACATCAAGAGCTTCGATATTAAAGGCTCAATGGCTTAAGCAGACAACAAATTT	1165
Db	1068	GGCTATGATGATGTGAAGAGATATGGAGACCAAGGTTTCAGTTCTTTAAAGAAATTTAAACCTG	1127
Qy	1166	GGAGGCCCATGGTCTGGGCCATTGATCTGGATGACTTCACTGGCACTTCTTGCAACCAG	1225
Db	1128	GGAGGACCATGATCTGGTCTATTGACATGGATGACTTCACTGGCAATCTCTGCAACCAG	1187
Qy	1226	GGCAAGTTTCCCCTAACTCCACCCTCGAAGAGGCCCTCGGC	1267
Db	1188	GGCCCTTACCTCTGTGCCAAGCAGTCAAGAGAGCCCTGGC	1229

RESULT 10

US-09-960-352-2589

; Sequence 2589, Application US/09960352

; Patent No. US20020137139A1

; GENERAL INFORMATION:

; APPLICANT: Warren, Wesley C.

; APPLICANT: Tao, Mengbing

; APPLICANT: Byatt, John C.

; APPLICANT: Mathialagan, Nagappan

; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND FAT DEPOSITION

; FILE REFERENCE: 16511.006/37-21(10298)C

; CURRENT APPLICATION NUMBER: US/09/960.352

; CURRENT FILING DATE: 2001-09-24

; NUMBER OF SEQ ID NOS: 15112

; SEQ ID NO 2589

; LENGTH: 415

; TYPE: DNA

; ORGANISM: Bos taurus

; OTHER INFORMATION: Clone ID: 12-LIB188-006-Q1-E1-C7

US-09-960-352-2589

Query Match 8.6%; Score 139.4; DB 10; Length 415;

Best Local Similarity 61.4%; Pred. No. 1.2e-35;

Matches 224; Conservative 0; Mismatches 141; Indels 0; Gaps 0;

Qy 187 CACCAACTGGGCCAGTACCGCCAGCGCTGGGGCGCTTCATGCCCTGACAACATCGACCC 246

Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 12 CGCATTGGGCCGGGATACCGGAGGTGATGGAGCTGCTTCCAGACGCCATCGACCC 71

Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 247 CTGCTCTGTACCCACCTGTATCAGCCCTTTGCTGGGAGGAGCAACAACGAGATCACCAC 306

Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 72 CTTCCTGTGCACCATGTCTATCTACAGCTTTGCCAACATAAGCAACAATGAGATCGACAC 131

Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 307 CATCGAATGAACGATGTGACTCTCTACCAAGCTTTCAATGGCCCTGAAAAATTAAGAACAG 366

Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 132 CTGGGAGTGAATGACGTGACGTCTATGACACACTGAACACACTCAAGAACAGGAACCC 191

Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 367 CCAGCTCAAAACTCTCTCGCCATTGGAGCTGGAACTTCGGGACTGCCCTTTCACTGC 426

Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 192 CAACCTGAAGACCTCTATCTGTGGAGATGGAATTCGGTCTCAAGATTTTCCAA 251

Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 427 CATGGTTTACTCTCTGAACGCCAGACTTTTCATCACCTCAGTCATCAAAATTCCTGG 486

Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 252 GATAGCTTCCAAGACCCAGAGTCGCAGGACTTTTCATCAAGTCGGTCCACCAATTTCTGG 311

Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 487 CCAGTATGAGTTTGAACGGCTGGACTTTGACTGGGAGTACCTCGTCTGGGGAGCC 546

Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 312 GACCATGGCTTTGATGAGCTGGACCTAGCATGGCTCTACCCCGGTGGAGACAAGCG 371

Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 547 TCCTC 551

Db 372 GCATC 376

RESULT 11

US-09-960-352-13050

; Sequence 13050, Application US/09960352

; Patent No. US20020137139A1

; GENERAL INFORMATION:

; APPLICANT: Warren, Wesley C.

; APPLICANT: Tao, Mengbing

; APPLICANT: Byatt, John C.

; APPLICANT: Mathialagan, Nagappan

; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND FAT DEPOSITION

; FILE REFERENCE: 16511.006/37-21(10298)C

; CURRENT APPLICATION NUMBER: US/09/960.352

; CURRENT FILING DATE: 2001-09-24

; NUMBER OF SEQ ID NOS: 15112

; SEQ ID NO 13050

; LENGTH: 414

; TYPE: DNA

ORGANISM: Bos taurus
OTHER INFORMATION: Clone ID: 56-LIB188-019-Q1-E1-F8
US-09-960-352-13050

Query Match 8.4%; Score 136.8; DB 10; Length 414;
Best Local Similarity 62.1%; Pred. No. 8.6e-35;
Matches 216; Conservative 0; Mismatches 132; Indels 0; Gaps 0;

QY 204 ACCGGCAGCGCTGGGGCTTCATCGCTGACAAATCGACCCCTCCCTGTGTGACCCACC 263
DB 1 ACCGGAGGGTGTATGGAGCTGCTCCAGAGCCCATGACCCCTCTGTGACCCCATG 60

QY 264 TGATCTACGCCCTTGTCTGGAGGAGCAGAACAGAGATCACCACTCATGATGAAAGCATG 323
DB 61 TCATCTACAGCTTTGCCAACAATAAGCAACAATGAGATCGACACCTGGGAGTGAATGACG 120

QY 324 TGACTCTACCAAGCTTTCAATGGCTGAAATTAAGACAGCCGAGCTGAAACCTCTCC 383
DB 121 TGACGCTCTATGACACATTTGAACACACTCAAGAACAGGAAACCCCAACCTGAGACCCCTCC 180

QY 384 TGGCCATTGGAGCTGGAACCTTGGGAGCTGCCCTTTCACTGCGCATGTTTCTACTCTCTG 443
DB 181 TATCTGTTGGAGATGGAACCTTGGTCTCAAGATTTTCCAGATAGCTTCCAAACACC 240

QY 444 AGAAGCGCAGACTTTATCACTACCTCATCAATTCCTGCGCCCATGATGATTTGACG 503
DB 241 GGAGTCCGAGGACTTTTCATCAAGTCGGTCCACCAATTTCTGGGACCCATGGCTTTGATG 300

QY 504 GCGTGCATTTGACTGGGAGTACCTGGCTCTCGTGGGAGCCCTCTCTC 551
DB 301 GACTGGACCTAGCATGGCTTACCCCGGGTGGAGACAAAGCGGCATC 348

RESULT 12

US-09-960-352-3072
Sequence 3072, Application US/09960352

Patent No. US20020137139A1
GENERAL INFORMATION:
APPLICANT: Warren, Wesley C.
APPLICANT: Tao, Nengbing
APPLICANT: Byatt, John C.
APPLICANT: Mathalagan, Nagappan
TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
FILE REFERENCE: 16511.006/37-21(10298)C
CURRENT APPLICATION NUMBER: US/09/960,352
CURRENT FILING DATE: 2001-09-24
NUMBER OF SEQ ID NOS: 15112
SEQ ID NO 3072
LENGTH: 418
TYPE: DNA
ORGANISM: Bos taurus
OTHER INFORMATION: Clone ID: 14-LIB188-015-Q1-E1-D5
US-09-960-352-3072

Query Match 8.4%; Score 136; DB 10; Length 418;
Best Local Similarity 60.3%; Pred. No. 1.6e-34;
Matches 251; Conservative 0; Mismatches 150; Indels 15; Gaps 1;

QY 203 TACCGCCAGCGCTGGGGCTTCATGCTTGACAAATCGACCCCTCCCTGTGTGACCCAC 262
DB 11 TACCGGAGGATGATGGAGCTGCTGCCAGAGCCCATGAGCCCTTCTGTGACCCCAT 70

QY 263 CTGATCTAGCCCTTCTCTGGGAGCAGAACAGGATCATCACCACATCGAATGGAACGAT 322
DB 71 GTCATCGACAGCTTTGCCACACATAAGCAACAATGATGATCGACACCTGGGAGTGAATGAC 130

QY 323 GTGACTCTTACCAAGCTTTCAATGGCTGAAATTAAGAACAGCCAGCTGAAACCTCTC 382
DB 131 GTGACGCTCTATGACACACTGACACACTCAAGAACAGGAAACCCCAACCTGAGACCCCTC 190

QY 383 CTGGCCATTTGGAGCTGGGAGTTCGAGCTGCCCCCTTCTACTGCCATGTTTCTACTCCT 442
DB 301 GACTGGACCTAGCATGGCTTACCCCGGGTGGAGACAAAGCGGCATC 348

Db 191 CTATCTGTTGGAGATGGAACCTTCGCTTCTCAAGATTTTCCAAAGATAGCTTCCAAAGACT 250

QY 443 GAGAACCGCAGACTTTTCATCCTCAGTCATCAATTCCTGCGCCAGTATGATTTGAC 502

Db 251 CAGAGTCGCGAGGACTTTTCATCAAGTCGGTGCCACCATTTCTGCGGACCCATGCTTTGAT 310

QY 503 GGGCTGGACTTTGACTGGGAGTACCTCGCTCTCGTGGGAGCCCTCTCTCAGGACAAAGCAT 562

Db 311 GGACTGGACTAGCATGGCTTACCCCGGGTGGAG-----AGACAAGCGG 355

QY 563 CTCTTCACCTGCTCTGCTGGAGAAATGCGTGAAGCTTTTCAGCAGGAGGCCAAGCA 618

Db 356 CATCTCACCACCTCTGCTCAAGGAAATCAAGGCTGAGTTTGTAAACGGAAGCCCAAGCA 411

RESULT 13
US-09-960-352-7364
Sequence 7364, Application US/09960352
Patent No. US20020137139A1
GENERAL INFORMATION:
APPLICANT: Warren, Wesley C.
APPLICANT: Tao, Nengbing
APPLICANT: Byatt, John C.
APPLICANT: Mathalagan, Nagappan
TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AN
FILE REFERENCE: 16511.006/37-21(10298)C
CURRENT APPLICATION NUMBER: US/09/960,352
CURRENT FILING DATE: 2001-09-24
NUMBER OF SEQ ID NOS: 15112
SEQ ID NO 7364
LENGTH: 410
TYPE: DNA
ORGANISM: Bos taurus
OTHER INFORMATION: Clone ID: 32-LIB188-023-Q1-E1-H1
US-09-960-352-7364

Query Match 8.3%; Score 135.2; DB 10; Length 410;
Best Local Similarity 61.8%; Pred. No. 2.9e-34;
Matches 215; Conservative 0; Mismatches 133; Indels 0; Gaps 0;

QY 204 ACCGGCAGCGCTGGGGCTTCATGCTGACAAATCGACCCCTCCCTGTGTGACCCACC 263
DB 1 ACCGGAGGGTGTATGGAGCTGCTCCAGAGCCCATGACCCCTTCTGTGACCCCATG 60

QY 264 TGATCTACGCCCTTGTCTGGGAGCAGAACAGGATCACCACTCGAATGGAACCATG 323
DB 61 TCATCTACAGCTTTGCCAACAATAAGCAACAATGAGATCGACACCTGGGAGTGAATGACG 120

QY 324 TGACTCTCTACCAAGCTTTCAATGGCTGAAATTAAGAACAGCCAGCTGAAACCTCTCC 383

Db 121 TGACGCTCTATGACACACTGACACACTCAAGAACAGGAAACCCCAACCTGAGACCCCTCC 180

QY 384 TGGCCATTGGAGCTGGAACCTCGGAGCTGCCCTTTCATGCGCATGTTTCTACTCTCTG 443

Db 181 TATCTGTTGGAGATGGAACCTCGGTTTCTCAAAGATTTTCCAAAGATAGCTTCCAAAGACC 240

QY 444 AGAAGCGCAGACTTTTCATCACCCTCATCAATTCCTGCGCCAGTATGATTTGACG 503

Db 241 GGAGTCCGAGGACTTTTCATCAAGTCGGTGCACCAATTTCTGCGGACCCATGGCTTTGATG 300

QY 504 GCGTGCATTTGACTGGGAGTACCTGCTGCTGGGAGCCCTCTCTC 551

Db 301 GACTGGACCTAGCATGGCTTACCCCGGGTGGAGACAAAGCGGCATC 348

RESULT 14
US-09-960-352-3057
Sequence 3057, Application US/09960352
Patent No. US20020137139A1
GENERAL INFORMATION:
APPLICANT: Warren, Wesley C.
APPLICANT: Tao, Nengbing

APPLICANT: Byatt, John C.
APPLICANT: Mathialagan, Nagappan
TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
FILE OF INVENTION: MUSCLE AND FAT DEPOSITION
FILE REFERENCE: 16511.006/37-21(10298)C
CURRENT APPLICATION NUMBER: US/09/960,352
CURRENT FILING DATE: 2001-09-24
NUMBER OF SEQ ID NOS: 15112
SEQ ID NO 3057
LENGTH: 424
TYPE: DNA
ORGANISM: Bos taurus
OTHER INFORMATION: Clone ID: 14-LIB188-002-Q1-E1-D5
US-09-960-352-3057

Query Match 8.3%; Score 135.2; DB 10; Length 424;
Best Local Similarity 61.8%; Pred. No. 3e-34;
Matches 215; Conservative 0; Mismatches 133; Indels 0; Gaps 0;
Qy 204 ACCGGCCAGGCTGGGGCGCTTCATGCTGACAACTACGACCCCTGCTGTACCCACC 263
Db 1 ACCGGGAGGTGATGGGAGCTGCTCCAGAGCCATCGACCCCTTCTGTGACCCCATG 60
Qy 264 TGATCTAGCGCTTGTGGGGAGGAGACACACAGAGATCACCATCATGATGGAAAGATG 323
Db 61 TCATCTACAGCTTGGCAACATAAGCAACATGAGATCGACACCTGGGAGTGGAAATGACG 120
Qy 324 TGACTCTCTACCAAGCTTTCATGGCTGAAATAAGAACAGCAGCTGAAACTCTCC 383
Db 121 TGAGCTCTATGACACTGNACACTCAAGACAGAACCCCAACTGGAAGACCTCC 180
Qy 384 TGGCCATGGAGGTGGAACCTGCGGAGTGGGACTGCCCTTTCACATGCTGCTTCTACTCTG 443
Db 181 TATCTGTGGAGGATGAACTTCGGTTCCTCAAGATTTTCCAGATAGCTTCCAGACCC 240
Qy 444 AGACCCGACAGCTTTCATCAGCTCAGTCAATTCCTCGCGCCAGTATGAGTTTGACG 503
Db 241 GGATCGCAGAGCTTTCATCAAGTCCGTGCGCACCATTCTCGGAGCCCATGGCTTTGATG 300
Qy 504 GGCTGGACTTGTACTGGGAGTACCTGGCTCTCGTGGAGCCCTCTCTC 551
Db 301 GACTGGACCTAGCATGCTCTACCCGGGTGGAGAGACAGCGGCATC 348

RESULT 15
US-09-960-352-14919
Sequence 14919, Application US/09960352
Patent No. US20020137139A1
GENERAL INFORMATION:
APPLICANT: Warren, Wesley C.
APPLICANT: Tao, Nengbing
APPLICANT: Byatt, John C.
APPLICANT: Mathialagan, Nagappan
TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
FILE OF INVENTION: MUSCLE AND FAT DEPOSITION
FILE REFERENCE: 16511.006/37-21(10298)C
CURRENT APPLICATION NUMBER: US/09/960,352
CURRENT FILING DATE: 2001-09-24
NUMBER OF SEQ ID NOS: 15112
SEQ ID NO 14919
LENGTH: 411
TYPE: DNA
ORGANISM: Bos taurus
FEATURE:
NAME/KEY: unsure
LOCATION: (4)
OTHER INFORMATION: unsure at all n locations
OTHER INFORMATION: Clone ID: 64-LIB188-023-Q1-E1-H12
US-09-960-352-14919

Query Match 7.5%; Score 122; DB 10; Length 411;
Best Local Similarity 61.8%; Pred. No. 7.2e-30;
Matches 194; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

Qy 238 CATGACCCCTGCTGTACCCACCTGATCTACGCCCTTGTGGGAGGAGCAACAAGCA 297
Db 36 CATGACCCCTGCTGTGACCCCATGTCTATCAGCTTGTCCCAACATAAGCAACATGA 95
Qy 298 GATCACCACCATCGAATGGAAAGATGTACTCTTACCAAGCTTTCAATGGCCTGAAAAA 357
Db 96 GATCGACACCTGGGAGTGGAAATGACGTGACGCTCTATGACACACTGAACACACTCAAGAA 155
Qy 358 TAAGAACAGCCAGCTGAAACTCTCTGGCCCATTTGGAGGCTGGAACTTCGGGACTGCCCC 417
Db 156 CAGGAACCCCAACCTGAAGACCCCTCTATCTGTTGGAGGATGGAACCTTCGGTTCTCAAAG 215
Qy 418 TTTCACTGCCATGTTTCTACTCTCTGAGAACCCGACAGCTTTCATCACTCAGTCATCAA 477
Db 216 ATTTTCCAAGATAGCTTCCCAAGACCCGAGTCCGAGGACTTTCATCAAGTCCGTGCCACC 275
Qy 478 ATTCTGCGCCAGTATGAGTTTACGGGCTGGAGCTTTGACTGGGAGTACCTGCTGCTCG 537
Db 276 ATTTCTGCGGACCATGGCTTTGATGGAGTGGACCTAGCATGGCTCTACCCCGGTGGAG 335
Qy 538 TGGGAGCCCTCTCTC 551
Db 336 AGACAAGCGGCATC 349

Search completed: July 3, 2003, 10:57:23
Job time : 287.278 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2003, 04:32:16 ; Search time 2436.98 Seconds
(without alignments)
10799.277 Million cell updates/sec

Title: US-10-004-219B-2
Perfect score: 1625
Sequence: 1 gcttccagtcgtggtga.....ctgcaataaaatcagcagtc 1625

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

1: em_estba:*
2: em_esthum:*
3: em_estinv:*
4: em_estm:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Score	Match Length	ID	Description
1	1047	64.4	1533	AK008633 Mus muscu
2	1042.2	64.1	1535	AK008757 Mus muscu
3	1018.4	62.7	1534	AK008650 Mus muscu
4	997.4	61.4	1530	AK007573 Mus muscu
5	772	47.5	957	B1760250 603045190
6	714.8	44.0	1105	BM923480 AGENCOURT

7	708.8	43.6	1038	13	BI517797	BI517797 603042049
8	681.8	42.0	892	14	BQ231098	BQ231098 AGENCOURT
9	653.6	40.2	872	14	BQ231064	BQ231064 AGENCOURT
10	649.8	40.0	838	12	BG867815	BG867815 602786336
11	641.2	39.5	824	13	BI518221	BI518221 603042049
12	630	38.8	855	12	BG872136	BG872136 602792974
13	629	38.7	837	12	BG868741	BG868741 602787570
14	626	38.5	847	12	BG866775	BG866775 602786605
15	623.2	38.4	809	12	BG866417	BG866417 602785352
16	622.4	38.3	836	12	BG869655	BG869655 602789716
17	620.2	38.2	1347	14	BM924663	BM924663 AGENCOURT
18	619.4	38.1	833	12	BG871337	BG871337 602790509
19	617.8	38.0	880	12	BG871468	BG871468 602790678
20	617.4	38.0	863	12	BG872850	BG872850 602793973
21	614.6	37.8	894	12	BG869958	BG869958 602789826
22	611.2	37.6	874	12	BG865756	BG865756 602784345
23	610.6	37.6	804	12	BG872805	BG872805 602793325
24	608.4	37.4	839	12	BG870376	BG870376 602791289
25	607.6	37.4	856	12	BG870147	BG870147 602793390
26	607.4	37.4	837	12	BG865578	BG865578 602783726
27	607.4	37.4	845	12	BG865848	BG865848 602788264
28	606.6	37.3	845	12	BG867145	BG867145 602786759
29	603.2	37.1	769	12	BG868463	BG868463 602785844
30	603	37.1	822	12	BG870367	BG870367 602791260
31	602.6	37.1	799	12	BG872144	BG872144 602792982
32	602	37.0	763	12	BG872647	BG872647 602793705
33	601.8	37.0	794	12	BG869492	BG869492 602789190
34	601.6	37.0	764	12	BG869304	BG869304 602788819
35	600	36.9	848	12	BG869143	BG869143 602789020
36	598.2	36.8	939	14	BQ219632	BQ219632 AGENCOURT
37	597.6	36.8	831	12	BG867708	BG867708 602786811
38	597.2	36.8	838	12	BG870392	BG870392 602791292
39	596.8	36.7	886	12	BG870840	BG870840 602791891
40	596.2	36.7	882	12	BG873488	BG873488 602791730
41	595.8	36.7	948	13	BI664452	BI664452 603290028
42	595.6	36.7	817	12	BG873398	BG873398 602794360
43	595.4	36.6	845	12	BG865076	BG865076 602784229
44	595.2	36.6	631	13	BI761276	BI761276 603044029
45	595	36.6	1155	12	BG867473	BG867473 602788579

ALIGNMENTS

RESULT 1	AK008633	1533 bp	mRNA	linear	HTC 19-JAN-2002
LOCUS	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:2200003E03:chitinase, acidic, full insert sequence.				
DEFINITION	AK008633				
ACCESSION	AK008633.1	GI:12842941			
VERSION	HTC; CAP trapper				
KEYWORDS	Mus musculus (strain:C57BL/6J) adult male stomach cDNA to mRNA, clone_lib:RIKEN full-length enriched mouse cDNA library				
SOURCE	clone:2200003E03.				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.				
AUTHORS	1				
TITLE	Carninci, P. and Hayashizaki, Y.				
JOURNAL	High-efficiency full-length cDNA cloning				
MEDLINE	Meth. Enzymol. 303, 19-44 (1999)				
PUBMED	99279253				
REFERENCE	10349636				
AUTHORS	2				
TITLE	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.				
JOURNAL	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
MEDLINE	Genome Res. 10 (10), 1617-1630 (2000)				
PUBMED	20499374				
REFERENCE	11042159				


```

|||||
491 GTGAAGAAATCGTGAAGCTTTTGAGCAGGAGGCTATTGAGAGCAACAGGCCAGACTG 550
QY ATGGTCATGCTGCGAGTACGCTGGCATCTCCATATATCCAGTCTGGCTATGATGATCC 597
Db ATGGTTACTGCTGCTAGCTGGTGGGATTTCCACATCCAGGCTGGTATGATGATCCCT 610
QY CAATCTGCACAGTACCTGGAGTACATCATGTGATGACCTAGCACTCCATGCTCCTGG 757
Db GAACCTTCAAGTACCTGGATTCATCCATGTCATGACATATGACCTCATGGCTCCTGG 670
QY GAGGCTACATGAGGAACAGCCCCCTTACAAATACCCGACATCCACCGGCGACAC 817
Db GAGGCTACATGAGGGAATAGTCTCTTTACAAATACCTTACTGAGACTGGTGAACAT 730
QY GCCTACCTCAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 877
Db GCCTACCTCAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 790
QY AAGCTCATCTGTTGATTCCTTACCTATGAGACACAACTTCTGAGCAACCCCTCAAC 937
Db AAGCTCATCTGTTGATTCCTTACCTATGAGACACAACTTCTGAGCAACCCCTCTGAT 850
QY ACTGGAATGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 997
Db AATGGAATGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 910
QY GGGATCTGGCTTACTAGCAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1057
Db GGGATCTGGCTTACTAGCAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 970
QY GATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1117
Db GATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1030
QY ATCAAGAGCTTCGATTAAGGCTCAATGCTTAAAGCTTAAAGCTTAAAGCTTAAAG 1177
Db ATCAAGAGCTTCGATTAAGGCTCAATGCTTAAAGCTTAAAGCTTAAAGCTTAAAG 1090
QY GTCTGGGCAATGATCTGATGATGATGATGATGATGATGATGATGATGATGATG 1237
Db GTCTGGGCAATGATCTGATGATGATGATGATGATGATGATGATGATGATGATG 1150
QY CTAATCTCCACCTCAGGAAGGCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1297
Db CTAATCTCCACCTCAGGAAGGCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1210
QY CTGACTTCTACTTTGAACAAAGCCCTTGSCATATCCATGAGTTCACACTCCTGAC 1210
Db CTGACTTCTACTTTGAACAAAGCCCTTGSCATATCCATGAGTTCACACTCCTGAC 1357
QY CAGCCCATGAGCCCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1357
Db CAGCCCATGAGCCCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1261
QY GTGCTTCCGAGCCAGTACTACTCTCC-----AGGAAGTGGAGTGGGGTGGGA 1417
Db GTGCTTCCGAGCCAGTACTACTCTCC-----AGGAAGTGGAGTGGGGTGGGA 1417
QY AGCTCTGGAGGAGCTCGGGAGGAGTGGATTCGTGCTGCTGCTGCTGCTGCTGCTGCT 1417
Db AGCTCTGGAGGAGCTCGGGAGGAGTGGATTCGTGCTGCTGCTGCTGCTGCTGCTGCT 1321
QY CCGTGGCAATTAACAGAAATGCTTCTGGCACTCGGCTGAATGAGTACCTACACGAG 1477
Db CCGTGGCAATTAACAGAAATGCTTCTGGCACTCGGCTGAATGAGTACCTACACGAG 1381
QY CCGTGGCAATTAACAGAAATGCTTCTGGCACTCGGCTGAATGAGTACCTACACGAG 1537
Db CCGTGGCAATTAACAGAAATGCTTCTGGCACTCGGCTGAATGAGTACCTACACGAG 1381
QY AACTGCCAGGCGGGCTGCTTTCGACACCACTGCTGATGCTGCTGCTGCTGCTGCTG 1537
Db AACTGCCAGGCGGGCTGCTTTCGACACCACTGCTGATGCTGCTGCTGCTGCTGCTG 1381
QY CATTCGAGGAGGCTGCTTTCGACACCACTGCTGATGCTGCTGCTGCTGCTGCTG 1441
Db CATTCGAGGAGGCTGCTTTCGACACCACTGCTGATGCTGCTGCTGCTGCTGCTG 1441
QY TGA 1540
Db TAA 1444
```

RESULT 2
AK008757
LOCUS
DEFINITION Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:2210020N23:chitinase, acidic, full insert sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE

AK008757.1 GI:12843148
HTC: CAP trapper.
Mus musculus (strain:C57BL/6J) adult male stomach cDNA to mRNA,
clone_lib:RIKEN full-length enriched mouse cDNA library
clone:2210020N23.

ORGANISM

Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE

1 Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636

AUTHORS

2 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159

TITLE

3 Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kitsumai,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861

JOURNAL

20530913

MEDLINE

11076861

PUBMED

11076861

REFERENCE

4 Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,
Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S.,
Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I.,
Saito,T., Okazaki,Y., Gojohori,T., Bono,H., Kasukawa,T., Saito,R.,
Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,
Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H.,
Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G.,
Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M.,
Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H.,
Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N.,
Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C.,
Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D.,
Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P.,
Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P.,
Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H.,
Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H.,
Tayo-Oka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L.,
Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohsaki,S.,
and Hayashizaki,Y.

TITLE

Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)

JOURNAL

21085660

MEDLINE

11217851

PUBMED

11217851

REFERENCE

5 (bases 1 to 1535)
Arakawa,T., Baldarelli,R., Bono,H., Brownstein,M., Bult,C.,
Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T.,
Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hirooka,T., Hori,F.,
Hume,D., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T.,
Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S.,
Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K.,
Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Quackenbush,J.,
Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D.,
Schriml,L., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yamanaka,I.,

161 TCTGCTTACCAGCTGACATCTACTTCCACCACTGGCCAGTACCGCGGAGCCTCTGGG 220
162 TCTGCTTACCAGCTGACATCTACTTCCACCACTGGCCAGTACCGCGGAGCCTCTGGG 220
74 TCTGCTTACCAGCTGACATCTACTTCCACCACTGGCCAGTACCGCGGAGCCTCTGGG 133
221 CGCTTACGCTGACACATCGACCCCTGCTCTGTACCACTGATCTAGCCTTTGCT 280
134 AGCTTCAAGCTGATGACATTAACCCCTGCTGTACTCACCTGATCTAGCCTTTGCT 193
281 GGGAGGAGAACACGAGATCACCACATGATGGAACATGATGATCTCTACCAAGCT 340
194 GGGAGGAGAACACGAGATCACCACATGATGGAACATGATGATCTCTCTCTAAAGCT 253
341 TTCAATGGCTGAAAATAAGAACAGCAGCTGAAAACCTCTCTGGCCATTTGGAGGCTGG 400
254 TTCAATGACATGAAACACGAGAACGAACTGAAACCTCTCTGGCAATTTGGAGGCTGG 313
401 AACTTCGGGAGTCCCTTTCACTGCGATGGTTTCACTCTCTGAGAACCCGACAGCTTC 460
314 AACTTTGGAAGTCTCTCTTTCACTTACCATGGTTTCCACTTCTCAGAACCCGACAGCTTC 373
461 ATCACTTCAGTCAATTAATTCCTGCGCAGTATGATGTTTGCAGCGGCTGAGCTTGCAGTGG 520
374 ATTAACCTCAGTCAATTAATTCCTGCGCAGTATGATGTTTGCAGCGGCTGAGCTTGCAGTGG 433
521 GAGTACCTGCTCTCTGCTGGAGCCCTCTCAGGACAAAGCATCTTCACTGCTCTGGTG 580
434 GAATACCCAGGCTCAGCTGGAGCCCTCTCAGGACAAAGCATCTTCACTGCTCTGGTG 493
581 CAGGAATGCTGAGCTTTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 640
494 AAGGAATGCTGAGCTTTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 553
641 GTCACTGCTCAGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 700
554 GTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 613
701 CTGTACAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 760
614 CTTTCTAAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 673
761 GGTACACTGAGAGACAGCCCTCTACAAATACCCGACTGACACCCGAGGAGGAGGAGGAGG 820
674 GGTACACTGAGAGACAGCCCTCTACAAATACCCGACTGACACCCGAGGAGGAGGAGGAGG 732
821 TACCTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 880
733 TACCTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 792
881 CTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 940
793 CTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 851
941 GGAATGCT 1000
852 GGAATGCT 911
1001 ATCTGGCTTACTACGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1060
912 TTCTGGGCTTACTACGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 971
1061 GCGCTCAGGAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1120
972 GCGCTCAGGAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1031
1121 AAGAGCTTCAATTAAGGCTTCAATTAAGGCTTCAATTAAGGCTTCAATTAAGGCTTCAATTAAGGCT 1180
1032 AAGAGCTTCAATTAAGGCTTCAATTAAGGCTTCAATTAAGGCTTCAATTAAGGCTTCAATTAAGGCT 1091
1181 TGGGCTTCAATTAAGGCTTCAATTAAGGCTTCAATTAAGGCTTCAATTAAGGCTTCAATTAAGGCT 1240
1092 TGGGCTTCAATTAAGGCTTCAATTAAGGCTTCAATTAAGGCTTCAATTAAGGCTTCAATTAAGGCT 1151
1241 ATCTCCACCTGAGAGGCTTCAATTAAGGCTTCAATTAAGGCTTCAATTAAGGCTTCAATTAAGGCT 1300

1152 ACTTCTACTTTGAACAAAGCCCTTGGCATATCCACTGAAGTTGCACAGCTCTCTGACGTG 1211
1301 CCATTGAGCAATACTGCTGCTCCAGTGGCAGCGGGAGCGGAGGAGTACGAGC 1360
1212 CCTTCCGAGCCAGTACTCTCTCTCC-----AGGAAGTGGAGTGGGGTGGGAGC 1262
1361 TCTGGAGGAGCTCGGAGGAGCAGTGGATTCTGCTGCTCAGAGCAACAGCGCTCTACCCC 1420
1263 TCCGAGGAGGAGCTCTGGAGCAGTGGATTCTGCTGCTCAGAGCAACAGGAGTGGCTCTACCCCT 1322
1421 GTGCAAAATACGAATGCTTCTTGGCACTGGTGAATGAGTACGTCACGACGAGCAAC 1480
1323 GTGCAATGACGAATGCTTCTTGGCACTGGTGAATGAGTACGTCACGAGCAGCAT 1382
1481 TGGCAGCGGCTGCTTCTGCTGACACAGCAGTGGATTCTGCTGCTCAGAGCAACAGGAGTGGGATTAACCTGA 1540
1383 TGTCAAGCAGGCTTGTGTTTGTATACCACTGTAATGCTGCACTGGCCATGAACCTAA 1442

RESULT 4
AK007573
LOCUS
DEFINITION
Mus musculus 10 day old male pancreas cDNA, RIKEN full-length
enriched library, clone:1810021D23:chitinase, acidic, full insert
sequence.
ACCESSION
AK007573
VERSION
AK007573.1 GI:12841201
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (Strain:C57BL/6J) 10 day old male pancreas cDNA to
mRNA, clone_lib:RIKEN full-length enriched mouse cDNA library
clone:1810021D23.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
MEDLINE
10349636
REFERENCE
2
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20493374
MEDLINE
11042159
REFERENCE
3
Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multipillar sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
MEDLINE
11076861
REFERENCE
4
Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,
Arakawa,T., Harai,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S.,
Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I.,
Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R.,
Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,
Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H.,
Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G.,
Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M.,
Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Anon,H.,
Baldarelli,R., Barsh,G., Blake,J., Botfelli,D., Bojunga,N.,


```
QY 1178 GTCTGGCCATGATCGGATGACTTCACCTGGGACCTTCCTGCAACAGGCGCAAGTTTCCC 1237
Db 1086 ATCTGGCCATGACCTTGATGACTTCACCTGGGCTCTTCTGTGATCAGGGAATTTCT 1145
QY 1238 CTATCTCCACCTGAAGAGGCCCTCGGCTCGAGTGCAGTTCAGCTCCAGCT 1297
Db 1146 CTGACTCTACTTTGAACAAGCCCTGGCATATCCACTGAAGTTGACAGCTCTCTGAC 1205
QY 1298 CAGCCATTTGAGCAATTAATCTGCTGCCAGTGGCAGCGGGAACGGAGGAGTAGC 1357
Db 1206 GTGCTTCGAGCCAGTACTCTCTCC-----AGGAAGTGGAGTGGGGTGA 1256
QY 1358 AGCTTGGAGGAGCTCGGAGGAGCTGATGATCTGCTGTCTAGAGCCCAACGCCCTCTAC 1417
Db 1257 AGCTCGGAGGAGCTCTGGAGGAGTGGATTCTGTGCCGACAAAGCAGATGCGCTCTAC 1316
QY 1418 CCCGTGGCAATTAACAGAAATGCTTCTGGCACTGGTGAATGGAGTCCAGTACCAGCAG 1477
Db 1317 CCGTGGCAGATGACAGAAATGCTTTTGGCAGTGCATCAATGGAATCAATACCAGCAG 1376
QY 1478 AACTGGCAGGCCGGCTGTCTTCGACACAGCTGTGATGCTGCAACTGGGATAAACC 1537
Db 1377 CATGTGACAGCGGCTGTTTGTATACCACTGTAAATGCTGCAACTGGCCATGACC 1436
QY 1538 TGA 1540
Db 1437 TAA.1439

RESULT 5
BI760250
LOCUS
DEFINITION BI760250 957 bp mRNA linear EST 25-SEP-2001
mRNA sequence.
BI760250
VERSION BI760250.1 GI:15751828
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 957)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11463 Row: c Column: 23
High quality sequence stop: 780.
Location/Qualifiers
1..957
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5185486"
/clone_lib="NIH_MGC_116"
/lab_host="DH10B"
/notes="Organ: pooled colon, kidney, stomach; Vector:
pCMV-SPORT6; Site:1: NotI; Site:2: EcoRV (destroyed); RNA
source anonymous pool of 3 colons, age 26 yo male, 49 yo
female, 71 yo male colon; 46 yo male kidney, and pool of 2
females, 62 yo male and 70 yo female. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.4 kb.
Insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
```

```
BASE COUNT 222 a 277 c 227 g 231 t
ORIGIN
Query Match 47.5%; Score 772; DB 13; Length 957;
Best Local Similarity 94.3%; Pred. No. 5.3e-204;
Matches 877; Conservative 0; Mismatches 45; Indels 8; Gaps 7;
339 CTTTCAATGCGCTGAAATAAAGACAGCAGCTGAAAACCTCTCTGGCCATTGGAGGCT 398
Db 25 CTTATCTAGATTTGACGCTCGGAACAGCAGCTGAAAACCTCTCTGGCCATTGGAGGCT 84
QY 399 GGAATTCGGAGCTGCCCTTTTCACTGCCATGTTTCTACTCTGAGAACCCGACACTT 458
Db 85 GGAATTCGGAGCTGCCCTTTTCACTGCCATGTTTCTACTCTGAGAACCCGACACTT 144
QY 459 TCATCACTCAGTCATCAAAATTCCTGGCCAGTATGATTTGACGGGCTGGACTTTGACT 518
Db 145 TCATCACTCAGTCATCAAAATTCCTGGCCAGTATGATTTGACGGGCTGGACTTTGACT 204
QY 519 GGGAGTACCTTGCTCTCGTGGAGGCCCTCTCAGGACAAGCATCTCTTCACTGTCTGG 578
Db 205 GGGAGTACCTTGCTCTCGTGGAGGCCCTCTCAGGACAAGCATCTCTTCACTGTCTGG 264
QY 579 TGCAGAAATGCGTGAAGCTTTTGGAGAGGAGGCCAAGCAGATCAACAAGCCAGGCTGA 638
Db 265 TGCAGAAATGCGTGAAGCTTTTGGAGAGGAGGCCAAGCAGATCAACAAGCCAGGCTGA 324
QY 639 TGGTCACTGCTGAGTAGCTGCTGCATCTCCAATATCCAGTCTGCATGATGATCCCC 698
Db 325 TGGTCACTGCTGAGTAGCTGCTGCATCTCCAATATCCAGTCTGCATGATGATCCCC 384
QY 699 AACTGTACACAGTACCTGGACTACATCCATGTCTATGACCTACGACCTCCATGGCTCTCTGG 758
Db 385 AACTGTACACAGTACCTGGACTACATCCATGTCTATGACCTACGACCTCCATGGCTCTCTGG 444
QY 759 AGGGCTACACTGGAGAGAACAGCCCTCTTACAATACCCGACTGACACCGGAGCAAGC 818
Db 445 AGGGCTACACTGGAGAGAACAGCCCTCTTACAATACCCGACTGACACCGGAGCAAGC 504
QY 819 CCTACCTCAATGTGGATTATGTCACTACTGGAAGACATGAGCACCAGCTGAGA 878
Db 505 CCTACCTCAATGTGGATTATGTCACTACTGGAAGACATGAGCACCAGCTGAGA 564
QY 879 AGCTCATCTGTGGATTCCCTTACCTATGGACACAATTCATCTCTGAGCAACCCCTCCAACA 938
Db 565 AGCTCATCTGTGGATTCCCTTACCTATGGACACAATTCATCTCTGAGCAACCCCTCCAACA 624
QY 939 CTGGAATTTGGTGGCCCACTCTGTGTGGTCTCTGCTGGCCCTATGCAAGAGTCTG 998
Db 625 CTGGAATTTGGTGGCCCACTCTGTGTGGTCTCTGCTGGCCCTATGCAAGAGTCTG 684
QY 999 GGATCTGGGCTTACTAGAGATCTGTACCTTCTCTGAA-AAATGGAGCCACTCAGGATGG 1057
Db 685 GGATCTGGGCTTACTAGAGATCTGTACCTTCTCTGAA-AAATGGAGCCACTCAGGATGG 744
QY 1058 GATGCCCCCTC-AGGAAGTGCCTTATGCCTA-TCAGGGCAATGTGGGTGGCTAT--GA 1113
Db 745 GATGCCCCCTCAGGAAGTGCCTTATGCCTATTCAGGGCAGACAGTGTGGTGGCTATGAC 804
QY 1114 CAACATCAAGAGC-TTCGATATTAAAGCTCAATGGC-TTAAGCACACAATAATTTGGAGGC 1171
Db 805 CAACATCAAGAGCTTTTCGATATTAAAGCTCAATGGCTTTTAGGCACACAACCAATTTGAGGC 864
QY 1172 GCCATGGTCTGGCCATTGATCT-GGATGACTTCACTGTCACCTTTCTGCAACACCGGCA 1230
Db 865 GCCATGGTCTGGCCATTGATCTGCGATTAACCTTCCACTGGGCACTTCTTGCACACCGGCA 924
QY 1231 GTTTCCCTTAATCTCCACCCTCGAAGAGC 1260
Db 925 GTTTCCCTTAATTTCCACCCTGTAACAGCG 954
```

```

RESULT 6
BM923480      1105 bp      mRNA      linear      EST 12-MAR-2002
LOCUS      AGENCOURT_6623879 NTH_MGC_116 Homo sapiens cDNA clone IMAGE:5759117
DEFINITION      5', mRNA sequence.
ACCESSION      BM923480
VERSION      BM923480.1 GI:19373859
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 1105).
AUTHORS      NIH-MGC http://mgc.nci.nih.gov/.
TITLES      National Institutes of Health, Mammalian Gene Collection (MGC).
JOURNALS      Unpublished (1999)
COMMENTS      Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-remail.nih.gov
              Tissue Procurement: Life Technologies, Inc.
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
              DNA Sequencing by: Agencourt Bioscience Corporation
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: LLAM12804 row: a column: 06
              High quality sequence stop: 668.
              Location/Qualifiers
                1..1105
                  /organism="Homo sapiens"
                  /db_xref="taxon:9606"
                  /clone="IMAGE:5759117"
                  /lab_host="NIH_MGC_116"
                  /lab_host="DH10B"
                  /note="Organ: pooled colon, kidney, stomach; Vector:
                  pcMV-SPORT6; Site.1: NotI; Site.2: EcoRV (destroyed); RNA
                  source anonymous pool of 3 colons, age 26 yo male, 49 yo
                  female, 71 yo male colon; 46 yo male kidney, and pool of 2
                  stomachs, 62 yo male and 70 yo female. Library is
                  oligo-dT primed and directionally cloned (EcoRV site is
                  destroyed upon cloning). Average insert size 1.4 kb,
                  insert size range 1-3 kb. Library is normalized and
                  enriched for full-length clones and was constructed by C.
                  Gruber (Invitrogen). Research Genetics tracking code
                  023. Note: this is a NIH_MGC Library."
BASE COUNT      270 a 318 c 285 g 231 t 1 others
ORIGIN
Query Match      44.0%; Score 714.8; DB 14; Length 1105;
Best Local Similarity 97.3%; Pred. No. 5.5e-188;
Matches 759; Conservative 0; Mismatches 17; Indels 4; Gaps 3;

QY 584 GAAATGCGTGAAGCTTTGAGCAGAGGCGCCAGCAGATCAACAGCCAGGCTGATGTC 643
      |||||
DB 121 GAAATGCGTGAAGCTTTGAGCAGAGGCGCCAGCAGATCAACAGCCAGGCTGATGTC 180
      |||||
QY 644 ACTGCTGCAGTAGCTGCTGGCATCTCCAAATATCCAGTCTGGCTATGAGATCCCCCAACTG 703
      |||||
DB 181 ACTGCTGCAGTAGCTGCTGGCATCTCCAAATATCCAGTCTGGCTATGAGATCCCCCAACTG 240
      |||||
QY 704 TCACAGTACCTGGACTACATCCATGTCATGACCTACGACCTCCATGGCTCTCTGGGAGGCG 763
      |||||
DB 241 TCACAGTACCTGGACTACATCCATGTCATGACCTACGACCTCCATGGCTCTCTGGGAGGCG 300
      |||||
QY 764 TACACTGGAGAACAGCCCCCTCTACAAATACCGACTGACACCGGACGACACGCTTAC 823
      |||||
DB 301 TACACTGGAGAACAGCCCCCTCTACAAATACCGACTGACACCGGACGACACGCTTAC 360
      |||||
QY 824 CTCATGTGGATTATGTCATGAATCTGAGAGGACATGAGGACACCTGAGAGCTC 883
      |||||
DB 361 CTCATGTGGATTATGTCATGAATCTGAGAGGACATGAGGACACCTGAGAGCTC 420
      |||||
QY 884 ATCGTTGGATTCCCTACCTATGACACAACTTTCATCCTGAGCAACCCCTCCAACTGGA 943
      |||||

Db 421 ATCGTTGGATTCCCTACCTATGAGCACAACTTCATCTTGAGCAACCCCTCCAACTGGA 480
      |||||
QY 944 ATTGTTGCCCCCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1003
      |||||
DB 481 ATTGTTGCCCCCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
      |||||
QY 1004 TGGGCTTACTAGCAGATCTGTACCTTCTCTGAAAATGAGGCACATCAGGATGGATGCC 1063
      |||||
DB 541 TGGGCTTACTAGCAGATCTGTACCTTCTCTGAAAATGAGGCACATCAGGATGGATGCC 600
      |||||
QY 1064 COTCAGGAAGTGCCTTATGCTCCTATCAGGCAATGTGTGGTGTGGCTATGACAACTCAAG 1123
      |||||
DB 601 COTCAGGAAGTGCCTTATGCTCCTATCAGGCAATGTGTGGTGTGGCTATGACAACTCAAG 660
      |||||
QY 1124 AGCTTCGATATTAAGGCTCAATGGCTTAAGCACACAAATTTGGAGGCGCATGCTCTGG 1183
      |||||
DB 661 AGCTTCGATATTAAGGCTCAATGGCTTAAGCACACAAATTTGGAGGCGCATGCTCTGG 720
      |||||
QY 1184 GCATTGATCTGGATGACTTCACTGGCACTTTCTCAACCCAGGCAAGTTTCCCTTAATC 1243
      |||||
DB 721 GCATTGATCTGGATGACTTCACTGGCACTTTCTCG-ACCAGGCAAGTTTCCCTTAATC 779
      |||||
QY 1244 TCCACCC--TGAAGAAGGCCCTCGGCCCTGCAGAGTGCAAGTTGCA-CGGCTCCAGCTCAG 1300
      |||||
DB 780 TCCACCCCTGAAAGAAGGCCCTCGGCCCTGCAGAGTGCAAGTTGCAAGCTCCAGCTCAG 839
      |||||
QY 1301 CCCATTGAGCCCAATTAAGTCTGCTCCAGTGCAGCGGACGGGAGGGAGTACGAGC 1360
      |||||
DB 840 CCCATTGAGCCCAATTAAGTCTGCTCCAGTGCAGCGGACGGGAGGGAGTACGAGTAAAC 899
      |||||

RESULT 7
BI517797      1038 bp      mRNA      linear      EST 29-AUG-2001
LOCUS      603042049F1 NTH_MGC_116 Homo sapiens cDNA clone IMAGE:5182357 5',
DEFINITION      mRNA sequence.
ACCESSION      BI517797
VERSION      BI517797.1 GI:15342589
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 1038)
AUTHORS      NIH-MGC http://mgc.nci.nih.gov/.
TITLES      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNALS      Unpublished (1999)
COMMENTS      Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-remail.nih.gov
              Tissue Procurement: Life Technologies, Inc.
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
              DNA Sequencing by: Incyte Genomics, Inc.
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: LLAM1455 row: a column: 14
              High quality sequence start: 3
              High quality sequence stop: 788.
              Location/Qualifiers
                1..1038
                  /organism="Homo sapiens"
                  /db_xref="taxon:9606"
                  /clone="IMAGE:5182357"
                  /lab_host="NIH_MGC_116"
                  /lab_host="DH10B"
                  /note="Organ: pooled colon, kidney, stomach; Vector:
                  pcMV-SPORT6; Site.1: NotI; Site.2: EcoRV (destroyed); RNA
                  source anonymous pool of 3 colons, age 26 yo male, 49 yo
                  female, 71 yo male colon; 46 yo male kidney, and pool of 2
                  stomachs, 62 yo male and 70 yo female. Library is
                  oligo-dT primed and directionally cloned (EcoRV site is

```

destroyed upon cloning). Average insert size 1.4 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
023. Note: this is a NIH_MGC Library."

BASE COUNT 273 a 275 c 293 g 197 t
ORIGIN

Query Match 43.6%; Score 708.8; DB 13; Length 1038;
Best Local Similarity 93.8%; Pred. No. 2.5e-186;
Matches 784; Conservative 0; Mismatches 42; Indels 10; Gaps 4;

```
QY 769 TGGAGAGAACAGCCCTCTACAAATACCCGACTGACACCGGAGCAACGCCCTACCTCAA 828
DB 1 TGGAGAGAACAGCCCTCTACAAATACCCGACTGACACCGGAGCAACGCCCTACCTCAA 60

QY 829 TGTGGATTATGTCATGAATCTGGAAGGACAAATGGAGCACACAGCTGAGAGCTCATCGT 888
DB 61 TGTGGATTATGTCATGAATCTGGAAGGACAAATGGAGCACACAGCTGAGAGCTCATCGT 120

QY 889 TGGATTCCCTTACCTATGGACACAACTTCATCTTGAGCAACCCCTCCAAACACTGGAATTGG 948
DB 121 TGGATTCCCTTACCTATGGAGCAACAACTTCATCTTGAGCAACCCCTCCAAACACTGGAATTGG 180

QY 949 TGCCCCACCTCTGGTCTGGTCTCTGGGCCCTATGCCAAGAGCTTGGGATCTGGGC 1008
DB 181 TGCCCCACCTCTGGTCTGGTCTCTGGGCCCTATGCCAAGAGCTTGGGATCTGGGC 240

QY 1009 TTACTAGGATCTGTACCTTCTGAAATATGAGGACCACTCAGGGATGGATGCCCTCA 1068
DB 241 TTACTAGGATCTGTACCTTCTGAAATATGAGGACCACTCAGGGATGGATGCCCTCA 300

QY 1069 GGAAGTGCCTTATGCTATCAGGGCAATGTGGTGGTATGACAAACATCAAGAGCTT 1128
DB 301 GGAAGTGCCTTATGCTATCAGGGCAATGTGGTGGTATGACAAACAGCTCAAGAGCTT 360

QY 1129 CGATATTAAGGCTCAATGGCTTAAGCACAAACAAATTTGGAGGCGCCATGTCTGGCCAT 1188
DB 361 CGATATTAAGGCTCAATGGCTTAAGCACAAACAAATTTGGAGGCGCCATGTCTGGCCAT 420

QY 1189 TGATCTGGATGACTTCACTGGCACTTCTGCAACAGGCAAGTCTTCCCTTATCTCCAC 1248
DB 421 TGATCTGGATGACTTCACTGGCACTTCTGCAACAGGCAAGTCTTCCCTTATCTCCAC 480

QY 1249 CCTGAAGAAGGCCCTCGGCTCGAGAGTGCAAGTTCACGGCTCCAGCTCAGGCCATTGA 1308
DB 481 CCTGAAGAAGGCCCTCGGCTCGAGAGTGCAAGTTCACGGCTCCAGCTCAGGCCATTGA 540

QY 1309 GCCAATACTGTCTCCAGTGGCAGCGGGAACCGGAGGAGTAGCAGCTCTGGAGG 1368
DB 541 GCCAATACTGTCTCCAGTGGCAGCGGGAACCGGAGGAGTAGCAGCTCTGGAGG 600

QY 1369 CAGCTCGGAGGAGTGGATTCT - GTGCTGTGACAGCAACGCCCTTACCCCTGGCA 1426
DB 601 CAGCTCGGAGGAGTGGATTCTGTGCTGGCAGAGCAACAGCTCTAACCCCTGGG 660

QY 1427 AATA - - - - -ACAGAAATGCTTCTGTGCACTGGCTGAATGAGTCACTACCCAG - -CAGAA 1479
DB 661 CAAATTAACAGAGATGCCTTCTGGGCATGCTGCTGAATGAGTCACTACCCAGAGAC 720

QY 1480 CTGCGAGGCGGCTTGT - CTTCGACACAGCTGTGATTGCTGCAACTGGGATAAACCT 1538
DB 721 TTGCGAGGCGGCTTGTCTTCGAGACAGCTGTGAATGCTGCAACTGGGATTAACCT 780

QY 1539 GACCTGGTATATCCCTAGAGTCCAGCTCTTTTGTCTAGGACATCTGCCCC 1594
DB 781 GACCTGGTATATCCCTAGAGTCCAGCTCTTTTGTCTAGGACATCTGGGCC 836
```

RESULT 8
BQ231098 892 bp mRNA linear EST 02-MAY-2002
LOCUS BQ231098 NCI_CGAP_St1 Mus musculus cDNA clone
DEFINITION AGENCOURT_7578138

IMAGE:6051580 5', mRNA sequence.

BQ231098

VERSION BQ231098.1 GI:20412498

KEYWORDS EST.

SOURCE house mouse.

ORGANISM

Mus musculus

REFERENCE

1 (bases 1 to 892)

NIH-MGC <http://mgc.nci.nih.gov/>.

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Tissue Procurement: Jeffrey E. Green, M.D.

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM13305 Row: k Column: 05

High quality sequence stop: 772.

Location/Qualifiers

1..892

/organism="Mus musculus"

/db_xref="taxon:10090"

/clone="IMAGE:6051580"

/lab_host="NCI_CGAP_St1"

/note="Organ: stomach; Vector: pCMV-SPORT6; Site_1: SalI;

Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.

Average insert size 1.77 kb. Library constructed by Life

Technologies. Note: this is a NCI_CGAP Library."

Technologies. Note: this is a NCI_CGAP Library."

Technologies. Note: this is a NCI_CGAP Library."

Technologies. Note: this is a NCI_CGAP Library."

Technologies. Note: this is a NCI_CGAP Library."

Technologies. Note: this is a NCI_CGAP Library."

Technologies. Note: this is a NCI_CGAP Library."

Technologies. Note: this is a NCI_CGAP Library."

Technologies. Note: this is a NCI_CGAP Library."

Technologies. Note: this is a NCI_CGAP Library."

Technologies. Note: this is a NCI_CGAP Library."

Technologies. Note: this is a NCI_CGAP Library."

Technologies. Note: this is a NCI_CGAP Library."

Technologies. Note: this is a NCI_CGAP Library."

Technologies. Note: this is a NCI_CGAP Library."

Technologies. Note: this is a NCI_CGAP Library."

Technologies. Note: this is a NCI_CGAP Library."

Technologies. Note: this is a NCI_CGAP Library."

Technologies. Note: this is a NCI_CGAP Library."

Technologies. Note: this is a NCI_CGAP Library."

Technologies. Note: this is a NCI_CGAP Library."

Technologies. Note: this is a NCI_CGAP Library."

Technologies. Note: this is a NCI_CGAP Library."

Technologies. Note: this is a NCI_CGAP Library."

Technologies. Note: this is a NCI_CGAP Library."

Technologies. Note: this is a NCI_CGAP Library."

Technologies. Note: this is a NCI_CGAP Library."

Technologies. Note: this is a NCI_CGAP Library."

Technologies. Note: this is a NCI_CGAP Library."

Technologies. Note: this is a NCI_CGAP Library."

Technologies. Note: this is a NCI_CGAP Library."

Technologies. Note: this is a NCI_CGAP Library."

Technologies. Note: this is a NCI_CGAP Library."

Technologies. Note: this is a NCI_CGAP Library."

Technologies. Note: this is a NCI_CGAP Library."

Technologies. Note: this is a NCI_CGAP Library."

Technologies. Note: this is a NCI_CGAP Library."

Technologies. Note: this is a NCI_CGAP Library."

Technologies. Note: this is a NCI_CGAP Library."

Technologies. Note: this is a NCI_CGAP Library."

Technologies. Note: this is a NCI_CGAP Library."

Technologies. Note: this is a NCI_CGAP Library."

Technologies. Note: this is a NCI_CGAP Library."

Technologies. Note: this is a NCI_CGAP Library."

Technologies. Note: this is a NCI_CGAP Library."

Technologies. Note: this is a NCI_CGAP Library."

Technologies. Note: this is a NCI_CGAP Library."

QY 612 CCAAGCAGATCAACAGCCAGGCTGATGCTCACTGCTGAGTAGTCTGGCATCTCCA 671
Db 481 CTATTGAGAGAACAGGCCACAGACGATGTTACTGCTGCTGATGCTGGGATTTCCA 540
QY 672 ATATCCAGTCTGGCTATGAGATCCCAACTGTACAGTACCTGGAGTACATCCATGTCA 731
Db 541 ACATCCAGGCTGGCTATGAGATCCCTGAACTTCTTAAGTACCTGGATTTTCATCCATGTCA 600
QY 732 TGACCTACCACTCCATGCTGCTGGAGGGTACACTGGAGAGACACCCCTCTACA 791
Db 601 TGACATATACCTCCATGCTGCTGGAGGGTACACTGGGAGATAGTCTCTTTACA 660
QY 792 AATACCGACTGACACCGGACGACGCTTACCTCAATGTGATATGTGATGAATCT 851
Db 661 AATACCTACTGAGACTGTGAGTACCTGCTTACCTCAATGTGATATGTGATGAATCT 720
QY 852 GGAAGACAAATGAGACACAGCTGAGAGCTCATGTTGGATTCCTACTATGGACACA 911
Db 721 GGAAGAACAAATGAGACACAGCTGAGAGCTCATGTTGGATTCCTACTATGGACACA 780
QY 912 ACTTATCTCTGAGCAACCCCTCCACACTGGAATGGTGGTCCCTGCTGGTGGTGC 971
Db 781 CTTTATCTCTGAGAAACCCCTCTGATATGGAATGGTGGTCCCTGCTGGTGGTGC 840
QY 972 CTGCTGGGCTCTATGCAAGGAGTCTGGGATCTGGGCTTACTACGA 1017
Db 841 CTGCTGGGCTCTATGCAAGGAGTCTGGGATCTGGGCTTACTACGA 886

RESULT 9
BQ231064
LOCUS AGENCOURT_7578230 NCI_CGAP_St1 Mus musculus cDNA clone EST 02-MAY-2002
IMAGE:6051490 5', mRNA sequence.
BQ231064
BQ231064.1 GI:20412464
EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 872)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM13305 row: g column: 11
High quality sequence stop: 724.
FEATURES
Location/Qualifiers
1..872
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:6051490"
/lab_host="NCI_CGAP_St1"
/note="Organ: stomach; Vector: pCMV-SPORT6; Site.1: Salt;
Site.2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.77 kb. Library constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 218 a 222 c 205 g 225 t 2 others
ORIGIN

Query Match 40.2%; Score 653.6; DB 14; Length 872;
Best Local Similarity 85.9%; Pred. No. 5.7e-171;
Matches 736; Conservative 0; Mismatches 120; Indels 1; Gaps 1;

QY 100 RACCATGACAAAGCTATTCTCTCCTCACAGGCTCTTGCTTATATGAAATTTGCAGCTCG 159
Db 1 AAGCATGGGCAAGCTACTTCTCGTCACAGGCTGGCTCTTCTGCTGAATGCTCAGCTGG 60
QY 160 CTCTGCTCTACCAAGCTGACATGCTTACCAACTGGGCCAGTACCGCCAGGCTGGG 219
Db 61 GTCTGCTCAACTGATATGCTATTTTCAACCACTGGGCCAGTATCGCCAGGCTGG 120
QY 220 GCGCTTCATGCTGTGACAAATCGACCCCTGCTCTGTATCCACCTGATCTACGCCCTTGC 279
Db 121 GAGCTTCAAGCCTGATGACATTAACCCCTGCTGTACTACCTGATCTATGCTTGC 180
QY 280 TGGGAGGAGAAACAAGAGATCACCACTCGAATGGAAGAGTGTGACTCTCTACCAAGC 339
Db 181 TGGGATGCAAGAACATGAGATCACCACTAGATAGTAATGATGTTACTCTCTATTAAGC 240
QY 340 TTTCAATGCGCTGAAATAAAGAACAGCCAGCTGAAAACTCTCTCGCCCATTTGAGGCTG 399
Db 241 TTTCAATGACTTTGAAAAACAGGAACAGCAACTGAAAAACCCCTCTGGCAATTTGAGGCTG 300
QY 400 GAATTCGGAGCTGCCCCCTTTTCTACTGCGATGGTTTCTACTCTGTGAGAACGCCAGACTTT 459
Db 301 GAATTTGGAAGCTGCTCTTCTACTACATGGTTTCCACTTCTCAGAACGCCAGACCTT 360
QY 460 CATCACTCAGTCATCAAAATTTCTGCGCCAGTATGATTTTGACGGGCTGAGCTTTGACTG 519
Db 361 CATCACTCAGTCATCAAAATTTCTGCGTCAGTATGGTTTGTGAGACTGGACCTGAGCTG 420
QY 520 GGAGTACCTGCTCTGCTGGAGCCCTCTCAGGACAAAGCATCTCTTCACTGCTCTGCT 579
Db 421 GGAATACCCAGGCTACGCTGGGAGCCCTCTCAGGACAAAGCATCTCTTCACTGCTCTGCT 480
QY 580 CGAGGAATTCGTTGAAGCTTTTGAGCAGGAGCCAGCAGATCAACAAGCCCGAGGCTGAT 639
Db 481 GAAGGAATTCGCTGAAAGCTTTTGAGCAGGAGGCTATTTGAGAGCAACAGCCCGAGACTGAT 540
QY 640 GGTCACTGCTGAGTAGCTGCTGCAATATCCAGTCTGGCTATGAGATCCCAAC 699
Db 541 GGTACTGCTGCTGAGTGGGATTTCCACATCCAGGCTGGCTATGAGATCCCTGA 600
QY 700 ACTGTCAAGTACCTGGAGTACATCCATGCTATGAGCTTCCAGCTCCATGGCTCTGGGA 759
Db 601 ACTTTCTAAGTACCTGGATTTTATCCATGCTATGAGTATGACCTCCATGGCTCTGGGA 660
QY 760 GGGTACACTGGAGAGAACAGCCCTCTACAATATCCGAGCTGACACCCGAGCAGCAGC 819
Db 661 NGGCTACACTGGGAGAAATAGTCTCTTTTACAAATACCTTACTGAGACTGGTAGCAATGC 720
QY 820 CTACCTCAATGTGGATTTATGATGAACTACTGGAAGGACAATGGAGCAGCAGCTGAGAA 879
Db 721 CTACCTCAATGTGGATTTATGATGAACTATTTGGAAGAACATGGAGCCCGAGCTGAGAA 780
QY 880 GCTCATCTGTTGGATTTCCCTTACCTATGAGACACA-ACCTTCATCTGAGCAACCCCTCAACA 938
Db 781 GCTCATCTGTTGGATTTCCAGAGATGGACACACCCCTTCTCTCTGAGAAACCCCTCTGATA 840
QY 939 CTGGAATTTGGTGGCCCC 955
Db 841 ATGGGAATTTGGTGGCCCC 857

RESULT 10
BQ867815
LOCUS 602786336F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4912544 5',
DEFINITION mRNA sequence.
ACCESSION BQ867815
VERSION BQ867815.1 GI:14218355
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 838)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM10816 row: g column: 09
High quality sequence stop: 824.
Location/Qualifiers
1. 838
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4912544"
/lab_host="NCI_CGAP_SG2"
/note="Organ: salivary gland; Vector: pCMV-SPORT6; Site: 1;
NotI; Site: 2; SalI; Cloned unidirectionally. Primer: Oligo
dT. Average insert size 1.3 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 209 a 211 c 198 g 220 t
ORIGIN
Query Match 40.0%; Score 649.8; DB 12; Length 838;
Best Local Similarity 87.0%; Pred. No. 6.3e-170;
Matches 714; Conservative 0; Mismatches 107; Indels 0; Gaps 0;
QY 137 CTTTACTGAATTGGACGTGGCTGCTGCTACAGCTGACATGCTACTTCCCAACTGG 196
DB 4 CTTTCTGTAATGCTCAGCTGGGCTGCTGCTACAACTGATGCTATTTCCCAACTGG 63
QY 197 GCCAGTACGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 256
DB 64 GCCAGTATGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 123
QY 257 ACCACCTGATCAGCGCTTGTGGGAGGAGCAACAGATCACCACCATCGAATGG 316
DB 124 ACTCACCTGATCTATGCTTGTGGGATGCGAACAATGAGATCACCACCATAGATGG 183
QY 317 AACGATGTGACTCTACCAAGCTTCAATGGCTGAAATTAAGACACCGACCTGAAA 376
DB 184 AATGATGTACTCTCTATAAGCTTCAATGACTTGAAACAGGAACTGAA 243
QY 377 ACTCTCTGGCCATTGGAGCTGGAATTCGGGAGTGGCCCTTTCTACTGCGATGGTTCT 436
DB 244 ACCCTCTGCAATGGAGCTGGAATTCGGGAGTGGCTCTTCTACTACATGGTTTC 303
QY 437 ACTCTGAGAACCGCAGACTTTCATCACTGATCATCAAAATTCCTGCGCCAGTATGAG 496
DB 304 ACTTCTCAGAACCGCAGACTTTCATCACTGATCATCAAAATTCCTGCGTCAATGGG 363
QY 497 TTTGACGGCTGACCTTGTACTGGGAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 556
DB 364 TTTGATGGAGTGGACCTGGGATACCCAGGCTCAGCTGGGAGCCCTCTCTCAGGAC 423
QY 557 AAGCATCTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 616
DB 424 AAGCATCTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 483
QY 617 CAGATCAACAGCCCGAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 676
DB 484 GAGAGCAACAGCCCGAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 543
QY 677 CAGCTGGCTATGAGATCCCGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 736

544 CAGGTGGCTATGAGATCCCTGAACTTTCTTAAGTACCTGGATTTCATCCATGTCATGACA 603
QY 737 TAGCACTCCATGGCTCTGGAGGCTACACTGGAGAACAGCCCTCTACAAATAC 796
DB 604 TAGCACTCCATGGCTCTGGAGGCTACACTGGAGAACAGCTCTCTTTTCAAAATAC 663
QY 797 CCGACTGACACCGGCGAGCAAGCTCTACCTCAATGTGGATTATGTGTAAGTACTGGAAG 856
DB 664 CTTACTGAGACTGGTAGCAATGCTACCTCAATGTGGATTATGTGTAAGTACTGGAAG 723
QY 857 GACAATGGAGCACAGCTGAGAGCTCATGCTGTTGGATTCCCTACTTATGACACACATTC 916
DB 724 AACAATGGAGCCCGAGCTGAGAGCTCATGTTTGGATTCCCGAGTTTGGACACACCTTC 783
QY 917 ATCTGTAGCAACCCCTCCAACTGGAATGGTGCCCCAC 957
DB 784 ATCTGTAGCAACCCCTCTGTAATGGATTGGTTGGTGGCCCTAC 824
RESULT 11
BI518221/c
LOCUS 824 bp mRNA linear EST 29-AUG-2001
DEFINITION 603042049T1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5182357 3',
mRNA sequence.
ACCESSION BI518221
VERSION BI518221.1 GI:15343013
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 824)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM1455 row: a column: 14
High quality sequence start: 24
High quality sequence stop: 746.
Location/Qualifiers
1. 824
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5182357"
/lab_host="NIH_MGC_116"
/note="Organ: pooled colon, kidney, stomach; Vector:
pCMV-SPORT6; Site: 1; NotI; Site: 2; EcoRV (destroyed); RNA
source anonymous pool of 3 colons, age 26 yo male, 49 yo
female, 71 yo male colon; 46 yo male kidney, and pool of 2
stomachs, 62 yo male and 70 yo female. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.4 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
023. Note: this is a NIH_MGC Library."
BASE COUNT 176 a 228 c 227 g 193 t
ORIGIN
Query Match 39.5%; Score 641.2; DB 13; Length 824;
Best Local Similarity 95.9%; Pred. No. 1.6e-167;
Matches 733; Conservative 0; Mismatches 23; Indels 8; Gaps 7;
QY 819 CTTACTCAATGGGATTATGTCATGAACT-ACTGGAGGACAATGGAGCACCAGCTGAG 877


```
QY 824 CTCAATGTGG-ATTATGTCATGAACCTA-CTGGAAGGACAAATGAGGACCAACCTGAGAAAGC 881
|||||
DB 721 CTCAATGTGGCATTATGTCATGAACCTATCTGGAAGAACAAATGAGGCCAAGCTGAGAAAGC 780
|||||
QY 882 TCATCGTGGATTCCTTACCTATGGAACAACTTCATCTCGACCAACCC 931
|||||
DB 781 TCATGTGGATTCCTGAGATGGAACACACCTTCATCTCGAGAAACCC 830
|||||

RESULT 13
BG868741 837 bp mRNA linear EST 29-MAY-2001
LOCUS 602787570F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4913800 5',
mRNA sequence.
ACCESSION BG868741
VERSION BG868741.1 GI:14219281
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE NIH-MGC http://mgc.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapb-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10819 row: k column: 17
High quality sequence stop: 836.
Location/Qualifiers
1. 837
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4913800"
/clone_lib="NCI_CGAP_SG2"
/lab_host="DH10B (Ti phage-resistant)"
/note="Organ: salivary gland; Vector: pCMV-SPORT6; Site: 1;
NotI; Site: 2; SalI; Cloned unidirectionally. Primer: Oligo
dT. Average insert size 1.3 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 212 a 211 c 197 g 217 t
ORIGIN
Query Match 38.7%; Score 629; DB 12; Length 837;
Best Local Similarity 84.5%; Pred. No. 4.1e-164;
Matches 707; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

QY 102 CCATGACAAGCTATTCTCCACAGGCTCTTCTTACTGTAATGTCAGCTGGCT 161
|||||
DB 1 CGATGCCCAAGCTACTCTCTGTCACAGGCTGGCTCTTCTGCTGAATGCTCAGCTGGGT 60
|||||
QY 162 CTGCTTACAGCTGACATGCTACTTCCACCACTGGCCAGTACCGCCAGGCTGGGC 221
|||||
DB 61 CTGCTTACCAATCTGATGATGCTATTTCACCACTGGCCAGTATCGCCAGGCTGGGA 120
|||||
QY 222 GCTTCATGCTGACAACTGACCCCTGCTGTATCCCACTGATCTACGCTTTGCTG 281
|||||
DB 121 GCTTCAAGCTGATGACATTAACCCCTGCTGTGACTCACTGATCTATGCTTTGCTG 180
|||||
QY 282 GGAGGCAACAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 341
|||||
DB 181 GGATGCAACAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
|||||
QY 342 TCATGGCTGAAATATAGACAGCCAGCTGAAACTCTCTGCGCATGAGGCTGGA 401
|||||
```

```
DB 241 TCAATGACTTGAAAAACAGGAACAGCAAACTGAAAAACCCCTCTGGCAATTTGSGAGGTGGA 300
QY 402 ACTTTGGGACTGCCCCCTTTCACTGCGCATGGTTTCTACTCTGAGAACCCGCGACACTTTCA 461
|||||
DB 301 ACTTTGGAAGTCTCTCTTTCACTACCATGTTTCCACTTCTCAGAACCCGCGACACTTCA 360
|||||
QY 462 TCACCTCAGTCATCAAAATTCCTCGCCAGTATGAGTTTGACGGCTGGAGCTTTTGACTGGG 521
|||||
DB 361 TTACCTCAGTCATCAAAATTTCTCGTCAAGTATGGTTTGTATGGTCTGAGCTGGACCTGAGTGG 420
|||||
QY 522 AGTACCTCGCTCTCGTGGAGGCCCTCTCTCAGACAGCAAGCATCTCTTCACTGTCTGCTGTC 581
|||||
DB 421 AATACCCAGGCTCAGCTGGAGGCCCTCTCTCAGACAGCAAGCATCTCTTCACTGTCTGTCG 480
|||||
QY 582 AGGAAATGCTGAAGCTTTTGACGAGGAGGCCCAAGCAGATCAACAAGCCCAAGCTGATGG 641
|||||
DB 481 AGGAAATGCTGAAGCTTTTGACGAGGAGGCTATTGAGACCAACACAGCCCAAGCTGATGG 540
|||||
QY 642 TCACCTGCTGCAGTGTGCTGTCATCTCCAATATCCAGTCTGGCTATGAGATCCCCCAAC 701
|||||
DB 541 TTACTGCTGCTGAGCTGGTGGATTTCCAAACATCCAGGCTGGCTATGAGATCCCTGAAC 600
|||||
QY 702 TGTCACTACCTGAGTACATCATGTCATGACCTACGACCTCCATGCTCTCTGAGG 761
|||||
DB 601 TTCTAAGTACCTGGATTTTCATCATGTCATGATGATGATGATGATGATGATGATGATGATG 660
|||||
QY 762 GCTACACTGGAGAGAACAGCCCTCTACAAATACCCGACTGACACCGGACGCAAGCCCT 821
|||||
DB 661 GCTACACTGGAGAGATAGTCTCTTTACAAATACCCCTACTGAGACTGGTATGAGATGCT 720
|||||
QY 822 ACCTCAATGTGGATTATGTCATGACTACTGGAAGGACAAATGAGGACCAAGCTGAGAAAGC 881
|||||
DB 721 ACCTCAATGTGGATTATGTCATGACTATTGGAAGAACAAATGAGGACCAAGCTGAGAAAGC 780
|||||
QY 882 TCATCTGCTGGATTCCTTACCTATGAGACAACTTCTCTGACCAACCCCTCCCAACA 938
|||||
DB 781 TCATTTGTTGATTTCCAGAGATGAGACACACCTTCAATCCTGAGAAACCCCTCGATA 837
|||||
```

RESULT 14
BG866775

LOCUS 602786605F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4912608 5',
mRNA sequence.

ACCESSION BG866775

VERSION BG866775.1 GI:14217315

KEYWORDS EST.

SOURCE house mouse.

ORGANISM

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 847)

AUTHORS NIH-MGC http://mgc.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapb-r@mail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM10816 row: j column: 01

High quality sequence stop: 774.

Location/Qualifiers

1. 847

/organism="Mus musculus"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone="IMAGE:4912608"

/clone_lib="NCI_CGAP_SG2"

FEATURES

source

RESULT 15
BG866417
LOCUS
 BC866417 809 bp mRNA linear EST 29-MAY-2001
DEFINITION
 G0278535ZFl NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4911421 5',
 mRNA sequence.
ACCESSION
 BG866417

KEYWORDS	house mouse.
SOURCE	Mus musculus
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 809)
REFERENCE	NIH-MGC http://mgc.ncl.nih.gov/ .
AUTHORS	National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE	Unpublished (1999)
JOURNAL	Contact: Robert Strausberg, Ph.D.
COMMENT	Email: cgapbs-r@mail.nih.gov Tissue Procurement: Jeffrey E. Green, M.D. CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM10813 row: h column: 14 High quality sequence stop: 804. Location/Qualifiers 1. .809 /organism="Mus musculus" /strain="FVB/N" /db_xref="taxon:10090" /clone="IMAGE:4911421" /lab_lib="NCI_CGAP_SG2" /lab_host="DH10B (T1 phage-resistant)" /note="Organ: salivary gland; Vector: pCMV-SPORT6; Site:1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dr. Average insert size 1.3 Kb. Constructed by Life Technologies. Note: this is a NCI_CGAP Library." BASE COUNT 193 a 200 c 208 g 207 t 1 others ORIGIN Query Match 38.4%; Score 623.2; DB 12; Length 809; Best Local Similarity 86.4%; Pred. No. 1.7e-162; Matches 699; Conservative 0; Mismatches 109; Indels 1; Gaps 1; QY 376 AACCTCTCTGGCCATTGGAGCGTGGAACTTCGGCACTCGCCCTTTTCACATGCCATGGTTTC 435 DB 1 AACCTCTCTGGCAATTTGGAGCGTGGAACTTTGGAAGTGTCTCTTTCACATACCAGTTTC 60 QY 436 TACTCTTGAGAACCGCCAGACATTTTCATCACTCACTCAGTCATCAAAATTCCTGGCCAGTATGA 495 DB 61 CACTTCTCAAGACCGCCAGACCTTCATTACCTCAGTCATCAAAATTTCTGGTCAGTATGG 120 QY 496 GTTTGACGGCTGGACTTTGACTGGGAGTACCTTGCTCTCGTGGGAGCCCTCCTCAGGA 555 DB 121 GTTTGATGACTGGACCTGGACTGGGAATACCCAGGCTCACGTGGGAGCCCTCCTCAGGA 180 QY 556 CAAGCATCTCTTTCACATGCTCTGCTGGCAGGAATGCGTGAAGCTTTTGGAGCAGAGGCCAA 615 DB 181 CAAGCATCTTTCACATGCTCTGTTGGAAGNAATGCGTGAAGCTTTTGGACGAGAGGCTAT 240 QY 616 GCAGATCAACAGCCCGAGGCTGATGGTTCATCTGCTCAGTAGTCTGGCATCTCCCAAT 675 DB 241 TGAGAGCAACAGCCCGAGAGCTGATGGTTCATCTGCTGCTGAGCTTTTGGACGAGAGGCTAT 300 QY 676 CCAGTCTGGGTATGAGATGCCCAACTGTCCACAGTACCTGGAGCTACATCCATGTCATGAC 735 DB 301 CCAGGTGGGTATGAGATCCCTGAACCTTTTAAATTAACCTGGATTTTCATCCATGTCATGAC 360 QY 736 CTACGACCTTCATGGCTCTCTGGGAGGCGCTACACTGGAGAGCAATACGCCCCCTCTACAAATA 795

Db 361 ATATGACCTCCATGGCTCCTGGGAGGGCTACACTGGGAGAGTAAGTCCTCTTTACAAATA 420

QY	796	CCGACTGACACCGGAGCAACCCCTACCTCAATGTGGATTATGTCACTAACTACTGGAA	855
Db	421	CCCTACTGAGACTGGTAGCAATCCCTACCTCAATGTGGATTATGTCACTAACTATTGGAA	480
QY	856	GGACATGGAGCACCAGCTGAGAGCTCATCGTTGGATTCCCTACCTATGGACACAACCTT	915
Db	481	GAACAATGGAGCCCCAGCTGAGAGCTCATTTGGATTCCCAAGATATGGACACACCTT	540
QY	916	CATCCTGAGCAACCCCTCCCAACACTGGAATTGCTGGCCCACTCTGGTCTGGTCTGC	975
Db	541	CATCCTGAGAAACCCCTCTGATTAATGGAATTGCTGGCCCTACCTCTGGTATGGCCCTGC	600
QY	976	TGGGCCCTATGCCAAGAGCTCTGGGATCTGGGCTTACTACGAGATCTGTACCTTCTCTGAA	1035
Db	601	TGGGCCCTATACCAGACAGGCTGGGTTCTGGGCCCTACTATGAGATNTGCACCTTTCTGAG	660
QY	1036	AAATGGAGCCACTCAGGGAT-GGGATGCCCTCAGGAAGTGCCTTATGCCCTATCAGGGCA	1094
Db	661	AAGTGGAGCCACTGAGGCTCTGGGGATGCCCTCCCAAGAAGTGCCTATGCCCTATAGGCCA	720
QY	1095	ATGCTGGGTTGGCTATGCAACATCAAGAGCTTCGATATTAAGGCTCAATGGCTTAAGC	1154
Db	721	ACGAGTGGCTTGCTATGACATTATCAAGAGCTTCAGTGTAAAGGTCAGTGGCTTAAGC	780
QY	1155	ACAACAAATTTGGAGCGCCATGGTCTGG	1183
Db	781	AGAACAAATTTGGAGGTGCCATGATCCGG	809

Search completed: July 3, 2003, 08:40:34
Job time : 2441.98 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2003, 02:31:01 ; Search time 4011.96 Seconds
(without alignments)
11062.367 Million cell updates/sec

Title: US-10-004-219b-3

Perfect score: 1525

Sequence: 1 atggcagactactctctgt.....taaaattgtagccaaaaca 1525

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vi.*

30: em_htg_hum.*

31: em_htg_inv.*

32: em_htg_other.*

33: em_htg_mus.*

34: em_htg_pln.*

35: em_htg_rod.*

36: em_htg_mam.*

37: em_htg_vrt.*

38: em_sy.*

39: em_htgo_hum.*

40: em_htgo_mus.*

41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1525	100.0	1530	10	AF290003	AF290003 Mus muscu
2	1520.2	99.7	1537	10	BC034548	BC034548 Mus muscu
3	1519.8	99.7	1538	10	BC011134	BC011134 Mus muscu
4	1517.2	99.5	1526	10	AF154571	AF154571 Mus muscu
5	1041	68.3	1625	9	AF290004	AF290004 Homo sapi
6	959.2	62.9	1529	4	AB051629	AB051629 Bos tauru
7	856	56.1	1354	9	AB025008	AB025008 Homo sapi
8	854	56.0	1369	6	AX405989	AX405989 Sequence
9	714.6	46.9	1525	10	M94584	M94584 Mus musculu
10	713.6	46.8	1506	10	D87757	D87757 Mus musculu
11	708.2	46.4	1209	10	AY049765	AY049765 Mus muscu
12	706.6	46.3	1209	10	AY065557	AY065557 Mus muscu
13	669	43.9	4250	9	AK098814	AK098814 Homo sapi
14	665	43.6	1188	9	AB025009	AB025009 Homo sapi
15	460	30.2	994	10	MMU58900	MMU58900 Mus musculu
16	430.4	28.2	1562	5	BJA345054	BJA345054 Bufo japo
17	426.8	28.0	1633	9	HSU29615	HSU29615 Human chit
18	426.8	28.0	1636	6	AR206041	AR206041 Sequence
19	426.8	28.0	1636	6	AR212138	AR212138 Sequence
20	426.8	28.0	1636	6	AX108750	AX108750 Sequence
21	426.8	28.0	1643	6	AR172107	AR172107 Sequence
22	426.8	28.0	1710	9	HSU62662	HSU62662 Homo sapien
23	426.8	28.0	1713	6	AR172108	AR172108 Sequence
24	425.2	27.9	1656	6	AR206042	AR206042 Sequence
25	425.2	27.9	1656	6	AR212139	AR212139 Sequence
26	425.2	27.9	1656	6	AX108752	AX108752 Sequence
27	371.4	24.4	2237	9	MMU87259	MMU87259 Macaca mula
28	368.2	24.1	2198	9	HSU09550	HSU09550 Human ovidu
29	365.2	23.9	2525	10	MUSOGP	D32137 Mouse mRNA
30	365	23.9	1406	9	AF484550	AF484550 Macaca ra
31	363.4	23.8	2228	9	BABEDQSG	M59903 Papio hamad
32	362	23.7	2504	6	E09046	E09046 cDNA encodi
33	355.6	23.3	2353	10	MAU15048	MAU15048 Mesocricetu
34	355.6	23.3	2387	10	HAMOGP	D32218 Hamster mRN
35	350.8	23.0	2366	6	E09047	E09047 cDNA encodi
36	350.4	23.0	233175	2	AC125461	AC125461 Mus muscu
37	343.2	22.5	1433	6	AR042834	AR042834 Sequence
38	343.2	22.5	1434	9	HSU58514	HSU58514 Human chiti
39	341.6	22.4	1418	9	HSU49825	HSU49825 Human YKL-3
40	341.6	22.4	1449	9	BC011460	BC011460 Homo sapi
41	340.4	22.3	1500	9	HSU58515	HSU58515 Human chiti
42	339.4	22.3	1994	4	BOV95BOGP	D16639 Bovine mRNA
43	339.4	22.3	1994	6	E09045	E09045 cDNA encodi
44	333.2	21.8	1599	9	AK055165	AK055165 Homo sapi
45	331.4	21.7	2034	4	OAU16719	UAU16719 Ovis aries

ALIGNMENTS

RESULT 1
AF290003
LOCUS AF290003
DEFINITION Mus musculus acidic mammalian Chitinase precursor, mRNA, complete cds.
ACCESSION AF290003
VERSION AF290003.1
KEYWORDS GI:12597290
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1530)
AUTHORS Boot,R.G., Blommaert,E.F., Swart,E., Chauharali-van der Vlugt,K., Bijl,N., Moe,C., Place,A. and Aerts,J.M.

TITLE Identification of a novel acidic mammalian chitinase distinct from chitotriosidase
J. Biol. Chem. 276 (9), 6770-6778 (2001) *Man*
MEDLINE 21125893
PUBMED 11085997
REFERENCE 2 (bases 1 to 1530)
AUTHORS Boot, R.G., Verhoeck, M., Swart, E. and Aerts, J.M.
TITLE Direct Submission
JOURNAL Submitted (26-JUL-2000) Dept. of Biochemistry, Academic Medical Center, University of Amsterdam, Meibergdreef 15, Amsterdam 1105 AZ, The Netherlands
FEATURES
source Location/Qualifiers
1. .1530
/organism="Mus musculus"
/strain="BALB/c"
/db_xref="taxon:10090"
/sex="male"
/tissue_type="lung"
/dev_stage="9-11 weeks"
3. .1424
/EC_number="3.2.1.14"
/note="chitinase precursor; AMCase"
/codon_start=1
/product="acidic mammalian chitinase precursor"
/protein_id="AAG60018.1"
/db_xref="GI:12597291"
/translation="MAKLIVLGLALLNLAQLGSAYNLICYFTNWAQYRPLGSLFKPD
DINPCLTHLIVAFAGMONNEITIEWNVLYKARNDLKRNKSLKTLIAIGWNG
TAPFTWVTSQNRQFETSVIKFLQYQDGLDWEIPGSRSPQDKHLFTLVK
EMREAEQEAIESNPRMLVMTAAVAGISNIQAGYEIPELSKYLDTHVMYDILHGSW
EGYTGNSPLKYPTGTGNSAYLVNDVYVWYKRNKGAPEKLIQVPEYGHFTILRNP
SDNGIGAPYSGDGPAGAYTRQAGWAYEICTFLRSAGATEVMDSEVPYAYKANLW
GYDNKISFVKAQWLKONNFGAMIAWLDLDFGFCDOGKPLTLNLKALGISTE
GCTAPDVPSPEPVTTPGSGSGSGSGSGSGSGGFCADKADGLYPVADDRNFWQCING
ITYQHQCAQLVDTSCNCCNWF"
BASE COUNT 382 a 383 c 366 g 399 t
ORIGIN
Query Match 100.0%; Score 1525; DB 10; Length 1530;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1525; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGCCCAAGTACTTCGCTGACAGGTCGGCTCTTCTGCTGAATGCTCAGCTGGGTCT 60
Db 3 ATGCCCAAGTACTTCGCTGACAGGTCGGCTCTTCTGCTGAATGCTCAGCTGGGTCT 62
Qy 61 GCCTACAACTGATATGCTATTTCACCACTGGCCCGCAGTATCGGCCAGCTCTGGGAGC 120
Db 63 GCCTACAACTGATATGCTATTTCACCACTGGCCCGCAGTATCGGCCAGCTCTGGGAGC 122
Qy 121 TTCAAGCCTGATGACATTAACCCCTGCTGTACTACCTGATGCTATGCTTTGCTGGG 180
Db 123 TTCAAGCCTGATGACATTAACCCCTGCTGTACTACCTGATGCTATGCTTTGCTGGG 182
Qy 181 ATGCAGAACATGAGATCACCACCATAGATGAATGATGTTACTCTCTATAAGCTTTC 240
Db 183 ATGCAGAACATGAGATCACCACCATAGATGAATGATGTTACTCTCTATAAGCTTTC 242
Qy 241 AATGACTTGAAGAACAGGAACAGCAACCTGCAATGCAATTTGGAGGCTGGAAC 300
Db 243 AATGACTTGAAGAACAGGAACAGCAACCTGCAATGCAATTTGGAGGCTGGAAC 302
Qy 301 TTTTGAAGTCTCTTTTCACTACCATGTTTCCACTTCTCAGAACCGCCAGACTTCATT 360
Db 303 TTTTGAAGTCTCTCTTTCACTACCATGTTTCCACTTCTCAGAACCGCCAGACTTCATT 362
Qy 361 ACCTCAGTATCAAAATTCGCTCAGTATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGG 420
Db 363 ACCTCAGTATCAAAATTCGCTCAGTATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGG 422
Qy 421 TACCCAGGCTCAGTGGGAGCCCTCTCAGGACAGCATCTTCTACTGCTCTGCTGCTGAAG 480
Db 423 TACCCAGGCTCAGTGGGAGCCCTCTCAGGACAGCATCTTCTACTGCTCTGCTGCTGAAG 482

Qy 481 GAAATGCGTGAAGCTTTTGAGCAGGAGGCTTATTGAGAGCAACAGGCCAGACTGATGTT 540
Db 483 GAATGCGTGAAGCTTTTGAGCAGGAGGCTTATTGAGAGCAACAGGCCAGACTGATGTT 542
Qy 541 ACTGCTGCTAGCTGGTGGGATTTCCAAACATCAGGCTGGCTATGAGATCCCTGAACCT 600
Db 543 ACTGCTGCTAGCTGGTGGGATTTCCAAACATCAGGCTGGCTATGAGATCCCTGAACCT 602
Qy 601 TCTAAGTACCTGATTTTCATCCATGTCATGACATATGACCTCCATGGCTCTGGGAGGC 660
Db 603 TCTAAGTACCTGATTTTCATCCATGTCATGACATATGACCTCCATGGCTCTGGGAGGC 662
Qy 661 TACATGGGGAGAAATAGTCTCTTTACAAATACCTACTGAGACTGGTAGCAATGCTCTAC 720
Db 663 TACATGGGGAGAAATAGTCTCTTTACAAATACCTACTGAGACTGGTAGCAATGCTCTAC 722
Qy 721 CTCATGTGGATTTATGTCATGAATATTTGGAAGAACAAATGGAGCCAGCTGAGAGCTC 780
Db 723 CTCATGTGGATTTATGTCATGAATATTTGGAAGAACAAATGGAGCCAGCTGAGAGCTC 782
Qy 781 ATTGTTGGATTTCCAGAGTATGACACACACCTTCTCTGAGAACCCCTCTGATATGGA 840
Db 783 ATTGTTGGATTTCCAGAGTATGACACACACCTTCTCTGAGAACCCCTCTGATATGGA 842
Qy 841 ATTGTTGGATTTCCAGAGTATGACACACACCTTCTCTGAGAACCCCTCTGATATGGA 900
Db 843 ATTGTTGGATTTCCAGAGTATGACACACACCTTCTCTGAGAACCCCTCTGATATGGA 902
Qy 901 TGGGCTCTACTATGAGATTTGCACCTTTCTGAGAAAGTGGAGCCACTGAGGCTGGGATGCC 960
Db 903 TGGGCTCTACTATGAGATTTGCACCTTTCTGAGAAAGTGGAGCCACTGAGGCTGGGATGCC 962
Qy 961 TCCCAAGAGTGGCTCTGCTATAGGCAACAGTGGCTTGGCTATGACAAATATCAAG 1020
Db 963 TCCCAAGAGTGGCTCTGCTATAGGCAACAGTGGCTTGGCTATGACAAATATCAAG 1022
Qy 1021 AGCTTCAGTGTAAAGGCTCAGTGGCTTAAAGCAGAACAAATTTTGGAGTGGCATGCTGG 1080
Db 1023 AGCTTCAGTGTAAAGGCTCAGTGGCTTAAAGCAGAACAAATTTTGGAGTGGCATGCTGG 1082
Qy 1081 GCCATTGACCTTGTATGACTTCACTGGCTCTTCTGATGATGAGGAAATTTTCTCTGACT 1140
Db 1083 GCCATTGACCTTGTATGACTTCACTGGCTCTTCTGATGATGAGGAAATTTTCTCTGACT 1142
Qy 1141 TCTACTTTGAAACAAAGCCCTTGGCATATCCACATGAAAGTTGACAGCTCTGACGTCCT 1200
Db 1143 TCTACTTTGAAACAAAGCCCTTGGCATATCCACATGAAAGTTGACAGCTCTGACGTCCT 1202
Qy 1201 TCCGAGCCAGTACTACTCTCCAGGAAGTGGAGTGGGGTGGAAAGCTCCGGAGGAAGC 1260
Db 1203 TCCGAGCCAGTACTACTCTCCAGGAAGTGGAGTGGGGTGGAAAGCTCCGGAGGAAGC 1262
Qy 1261 TCTGAGGAGTGGATTTCTGTCGAGAACAGAGATGGCTCTTACCTCTGGGAGAGTAC 1320
Db 1263 TCTGAGGAGTGGATTTCTGTCGAGAACAGAGATGGCTCTTACCTCTGGGAGAGTAC 1322
Qy 1321 AGAAATGCTTTTGGCAGTGCATCAATGGAATCACATACACAGCAGCATTTGCAAGCAGG 1380
Db 1323 AGAAATGCTTTTGGCAGTGCATCAATGGAATCACATACACAGCAGCATTTGCAAGCAGG 1382
Qy 1381 CTTGTTTTGATACAGCTGTAAATGCTGCAATGGCCATGAAACCTTAATGCCATTTCTTC 1440
Db 1383 CTTGTTTTGATACAGCTGTAAATGCTGCAATGGCCATGAAACCTTAATGCCATTTCTTC 1442
Qy 1441 AGAAATTTCTGCACTCTCTTTTACTCTCACCAAAAGTAACTATCTCCCTTTAACCTTA 1500
Db 1443 AGAAATTTCTGCACTCTCTTTTACTCTCACCAAAAGTAACTATCTCCCTTTAACCTTA 1502
Qy 1501 TGCAATAAAATTTGGTAGCCAAAACA 1525
Db 1503 TGCAATAAAATTTGGTAGCCAAAACA 1527

RESULT 2	BC034548	linear	1557 bp	MRNA	ROD 07-AUG-2002							
LOCUS	BC034548											
DEFINITION	mus musculus, chitinase, acidic, clone MGC:15045 IMAGE:4189080, mRNA, complete cds.											
ACCESSION	BC034548											
VERSION	BC034548.1	GI:21961190										
KEYWORDS	MGC.											
SOURCE	house mouse.											
ORGANISM	mus musculus											
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.											
AUTHORS	1 (bases 1 to 1557)											
TITLE	Strausberg R.											
JOURNAL	Direct Submission											
	Submitted (24-JUL-2002) National Institutes of Health, Mammalian											
	Gene Collection (MGC), Cancer Genomics Office, National Cancer											
	Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,											
	USA											
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov											
COMMENT	Contact: MGC help desk											
	Email: cgapbs-r@mail.nih.gov											
	Tissue Procurement: Jeffrey E. Green, M.D.											
	cDNA Library Preparation: The I.M.A.G.E. Consortium (LLNL)											
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)											
	DNA Sequencing by: Baylor College of Medicine Human Genome											
	Sequencing Center											
	Center code: BCM-HGSC											
	Web site: http://www.hgsc.bcm.tmc.edu/cdna/											
	Contact: amg@bcm.tmc.edu											
	Gunnarathne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,											
	Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,											
	Richards, S., Gibbs, R.A.											

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAK File: 24 Row: h Column: 16
 This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein.

FEATURES

```

1. .133/
/organism="Mus musculus"
/db_xref="taxon:10090"
/map="FVB/N"
/clone=WGC:19045 IMAGE:4189080"
/tissue_type="Salivary gland, 10 week old female mouse"
/clone_lib="NCI_CGAP_SG2"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
4. .1425
/codon_start=1
/product="chitinase, acidic"
/protein_id="AAH34548.1"
/db_xref="GI:21961191"
/db_xref="LocusID:81600"
/translation="WAKLLIVTGLALLNLAQLGSAYNLICYFTNWQRYRPLGSLGSKPFD
DINPCLTHLIATAFAGMONEITIEWNDVTLYKAFDLNNRNKSUKTLAIGGNWFNG
TAPPTTWTSQNQTSTVTKIFRLQYRFQGLDLWEYFGSGRSQPDKHLFTFLVK
EMEAPEAQEASISNRPLMTVAAGAAGLSIQAGCYEIPELSKYLDFTHVTVDLHGSK
GYTGNSPLKYKPYETGSNAVLNDVYMVNKKNGAPAEKLILVGPEPGYCHTFILRP
SDNICAPNTSGDGAPGVTRQANGFWAYIEICTFRSCATEKWDAQBVPYAYKANEWL
GYDNIKSFVKAQWLKNQNFGAALAILDDDTFGSCDGOKGPLETSTLNKAIGISTE
ITVAPQVDSBPVTPPGSSGGSSGGSSGGSSGGSKADKADGLYPVADRRAEFQQING
ITVQQHCQALVFDTSCACCNWP"
382 C 366 g 399 t

```

Query Match	99.7%	Score 1520.2;	DB 10;	Length 1557;
Best Local Similarity	99.8%	Pred. No. 0;		
Matches 1522;	Conservative	0; Mismatches	3; Indels	0; Caps
				0;

Qy	1	ATGCCAAGCTACTTCTCGTCACAGGTCTTGGCTCTTCTGCTGAATGCTCAGCTGGGGTCT	60
Db	4	ATGCCAAGCTACTTCTCGTCACAGGTCTTGGCTCTTCTGCTGAATGCTCAGCTGGGGTCT	63
Qy	61	GCCTACAACTGTGATATGCTATTTTACCAAACTGGGCCAGTATCGGCCAGAGCTCGGGGAGC	120
Db	64	GCCTACAACTGTGATATGCTATTTTACCAAACTGGGCCAGTATCGGCCAGAGCTCGGGGAGC	123
Qy	121	TTCAAGCGCTGATGACATTAACCCCTGGCTGTGTACTCACCTGATCTATGCGCTTTGCTGGG	180
Db	124	TTCAAGCGCTGATGACATTAACCCCTGGCTGTGTACTCACCTGATCTATGCGCTTTGCTGGG	183
Qy	181	ATGCAGAACAAATGAGATCACACATAGAATGAATGATGTTTACTCTCTATAAAGCTTTC	240
Db	184	ATGCAGAACAAATGAGATCACACATAGAATGAATGATGTTTACTCTCTATAAAGCTTTC	243
Qy	241	AATGACTTGAAACAGGAACACCAACTGAAACCCCTCGTGCATTTGGAGCTGGAAC	300
Db	244	AATGACTTGAAACAGGAACAGCAAACTGAAACCCCTCGTGCATTTGGAGCTGGAAC	303
Qy	301	TTTGGAACTGCTCCTTTTCACTACCATGGTTTCCACTTCTCAGAACCGCCAGACCTTCATT	360
Db	304	TTTGGAACTGCTCCTTTTCACTACCATGGTTTCCACTTCTCAGAACCGCCAGACCTTCATT	363
Qy	361	ACCTCAGTCATCAAAATTTCTGGTCACTATGCGTGTGATGGACTGGACCTGGAGAA	420
Db	364	ACCTCAGTCATCAAAATTTCTGGTCACTATGCGTGTGATGGACTGGACCTGGAGAA	423
Qy	421	TACCCAGGCTCAGTGGAGCCCTCCTCAGSACAAGCATCTTTCAGTGTCTCTGTTGTAAG	480
Db	424	TACCCAGGCTCAGTGGAGCCCTCCTCAGSACAAGCATCTTTCAGTGTCTCTGTTGTAAG	483
Qy	481	GAATGCGTGAAGCTTTTGAAGCAGGAGGCTATTGAGAGCAACAGGCCAGACCTGATGGTT	540
Db	484	GAATGCGTGAAGCTTTTGAAGCAGGAGGCTATTGAGAGCAACAGGCCAGACCTGATGGTT	543
Qy	541	ACTGCTGCTGATGTGGTGGATTTTCAACATCCAGGCTGGCTATGAGATCCCTGAACTT	600
Db	544	ACTGCTGCTGATGTGGTGGATTTTCAACATCCAGGCTGGCTATGAGATCCCTGAACTT	603
Qy	601	TCTAAGTACCTGGATTTTCATCCATGTCATCAGATATGACCTCCATGGCTCCTCGGAGGC	660
Db	604	TCTAAGTACCTGGATTTTCATCCATGTCATCAGATATGACCTCCATGGCTCCTCGGAGGC	663
Qy	661	TACACTGGGAGAAATAGTCCCTCTTTACAAATACCCCTACTCAGACTGGTATGCAATGCTTAC	720
Db	664	TACACTGGGAGAAATAGTCCCTCTTTACAAATACCCCTACTCAGACTGGTATGCAATGCTTAC	723
Qy	721	CTCAATGTGGATTATGTCATGAATATTGGAAGAACAAATGGAGCCCCAGCTGAGAAGCTC	780
Db	724	CTCAATGTGGATTATGTCATGAATATTGGAAGAACAAATGGAGCCCCAGCTGAGAAGCTC	783
Qy	781	ATTGTTGGATTCCAGAGTATGACACACACTTTCATCGTGAAGAACCCCTCTGATATATGGA	840
Db	784	ATTGTTGGATTCCAGAGTATGACACACACTTTCATCGTGAAGAACCCCTCTGATATATGGA	843
Qy	841	ATTGGTGCCCTTACCTCTGTGTGATGGCCCTGCTGGCCCTPATACACAGAGGCTGGGTTTC	900
Db	844	ATTGGTGCCCTTACCTCTGTGTGATGGCCCTGCTGGCCCTPATACACAGAGGCTGGGTTTC	903
Qy	901	TGGGCCCTACTATGAGATTTTGACCTTCTTGAGAAGTGGAGCCACTCAGGCTCTGGGATGCC	960
Db	904	TGGGCCCTACTATGAGATTTTGACCTTCTTGAGAAGTGGAGCCACTCAGGCTCTGGGATGCC	963
Qy	961	TCCCAAGAAGTGGCCCTATGGCTATAGGCCCAAGAGTGGCTTGGCTATGACAATATCAAG	1020
Db	964	TCCCAAGAAGTGGCCCTATGGCTATAGGCCCAAGAGTGGCTTGGCTATGACAATATCAAG	1023
Qy	1021	AGCTTCAGTCTTAAGGCTCAGTGGCTTAAAGCAACAATTTTGGAGGTCCTCATGATCTGG	1080
Db	1024	AGCTTCAGTCTTAAGGCTCAGTGGCTTAAAGCAACAATTTTGGAGGTCCTCATGATCTGG	1083
Qy	1081	GCCATTGACCTTGATGACTTCACTGGCTCTTTTCTGTGATCAGGGAATAATTTCTCTGACT	1140

[illegible]

```
QY 121 TTCAAGCCTGATGACATTAAACCCCTGCCTGTGTACTCACTGATCTATGCTCTTCTGGG 180
Db 224 TTATGCTGCTGACAAATGACCCCTGCCTCTGTACCCACCTGATCTAGCCCTTTCTGGG 283
QY 181 ATGCAGAACTGAGATCACCACCATAGAAATGAATGATGTACTCTCTATTAAGACTTTC 240
Db 284 AGGCAGAACTGAGATCACCACCATAGAAATGAATGATGTACTCTCTATTAAGACTTTC 343
QY 241 AATGACTTGAATAAAGGAGCAACAGCAACCTTCCCTGGCAATTTGGAGCTTGGAAAC 300
Db 344 AATGCTCTGAATAAATAGAACAGCAGCTGAAATCTCTCTGGCCATTTGGAGCTTGGAAAC 403
QY 301 TTGGAAGCTGCTCTTCACTACCATGTTTCCACTTCTCAGAACCCGACAGACTTCAAT 360
Db 404 TTCGGGACTGCCCCCTTCACTGCCATGTTTCTACTCTCTGAGAACCGCAGACTTTCATC 463
QY 361 ACCTCAGTCACTAAATTTCTGGCTCAGTATGCTGTTGATGGACTGGACCTGGACTGGGNA 420
Db 464 ACCTCAGTCACTAAATTTCTGGCCAGTATGATTTGACGGCTGGACTTTGACTGGGAG 523
QY 421 TACCAGGCTCAGCTGGGAGCCCTCTCAGGACAAAGCATCTCTTCACTGTCTCTGTGTAAG 480
Db 524 TACCCTGGCTCTCTGGGAGCCCTCTCAGGACAAAGCATCTCTTCACTGTCTCTGTGTAAG 583
QY 481 GAAATGCTGAGCTTTTGGAGAGGCTATTTGAGAGCAACAGCCACAGCTGATGTT 540
Db 584 GAAATGCTGAGCTTTTGGAGAGGCTATTTGAGAGCAACAGCTGATGTT 643
QY 541 ACTGCTGCTGAGCTTGGTGGATTTCCAACTCCAGGCTGGCTATGAGATCCCTGAATTT 600
Db 644 ACTGCTGAGTATGCTGGCATCTCCAAATATCCAGTCTGGTATGAGATCCCTGAATTT 703
QY 601 TCTAATGCTGATCTGATCATGCTGATGATGATGATGATGATGATGATGATGATGATGAT 660
Db 704 TCACAGTACCTGGACTATCATGCTGATGATGATGATGATGATGATGATGATGATGATGAT 763
QY 661 TACACTGGGAGAAATGATCTCTTTTACAAATACCTTACTGAGACTGGTAGCAATGCTAC 720
Db 764 TACACTGGGAGAAATGATCTCTTTTACAAATACCTTACTGAGACTGGTAGCAATGCTAC 823
QY 721 CTCATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
Db 824 CTCATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 883
QY 781 ATTTGTTGATTTCCAGAGATGACACACTTTCATCTCTGAGAAACCCCTCTGATATGA 840
Db 884 ATCTGTTGATTTCCATACCTATGACACACTTTCATCTCTGAGAAACCCCTCTGATATGA 943
QY 841 ATTTGTTGCTTACCTCTGTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Db 944 ATTTGTTGCTTACCTCTGTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1003
QY 901 TGGGCTTACTATGATTTGACCTTCTGAGAAAGTGGAGCCACTGAGTCTGGGATGCC 960
Db 1004 TGGGCTTACTATGATTTGACCTTCTGAGAAAGTGGAGCCACTGAGTCTGGGATGCC 1063
QY 961 TCCCAAGAGTGCCTATGCTTAAAGGCAACAGTGGCTTGGCTATGACAAATATCAAG 1020
Db 1064 CCTCAGGAAGTGCCTATGCTTAAAGGCAACAGTGGCTTGGCTATGACAAATATCAAG 1123
QY 1021 AGCTTCAGTGTAAAGCTCAGTGGTAAAGCAGAAATTTTGGAGGTGCCATGATCTGG 1080
Db 1124 AGCTTCAGTGTAAAGCTCAGTGGTAAAGCAGAAATTTTGGAGGTGCCATGATCTGG 1183
QY 1081 GCCATTGACCTTATGACTTCACTGCTCTTCTGATGATGATGATGATGATGATGATGATGAT 1140
Db 1184 GCCATTGATCTGGATGACTTCACTGCTCTTCTGATGATGATGATGATGATGATGATGATGAT 1243
QY 1141 TCTATTGTAACAAAGCCCTTGGCATATCCATGAGGTTCACACAGCTCTCTCAGCTGCTCT 1200
Db 1244 TCCACCTGTAACAAAGCCCTTGGCATATCCATGAGGTTCACACAGCTCTCTCAGCTGCTCT 1303
QY 1201 TCCGAGCCAGTACTACTCTCTCC-----AGGAAGTGGAGTGGGGGTGGAAGCTCC 1251
```

```
Db 1304 ATTGAGCCAATAACTGCTGCTCCAGTGGCAGCGGAACGGGAGGAGTAGACAGCTCT 1363
QY 1252 GGAGGAAGCTCTGGAGGAGTCTGCTGCGGCAACAGCAGATGGCTCTTACCTGG 1311
Db 1364 GGAGGAGCTCGGGAGGAGTCTGCTGCTGTCAGAGCAACGGCTCTTACCCGG 1423
QY 1312 CGAGATGACAGAAATGCTTTTGGGAGTGCATCAATGGAATCACATACCAAGCAGCATTTCT 1371
Db 1424 GCAATACAGAAATGCTTCTGGCAGTGCCTGGAATGAGTCACTACCAAGCAGACTGC 1483
QY 1372 CAAGCAGGCTTGTGTTTGTATACAGCTGTAATGCTGCAATGGCCTGCACTAA 1428
Db 1484 CAGCGCGGCTGCTGCTGACACAGCTGATGCTGCACTGGGCAATAAAGCTGA 1540

RESULT 6
AB051629
LOCUS
DEFINITION
  Bos taurus cbp b04 mRNA for chitin binding protein b04, complete
  cds.
ACCESSION
  AB051629
VERSION
  AB051629.1 GI:16754894
KEYWORDS
  Bos taurus liver cDNA to mRNA.
ORGANISM
  Bos taurus
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
  Bovidae; Bovinae; Bos.
REFERENCE
  1
  AUTHORS
    Suzuki,M., Morimatsu,M., Yamashita,T., Iwanaga,T. and Syuto,B.
  TITLE
    A novel serum chitinase that is expressed in bovine liver
  JOURNAL
    FEBS Lett. 506 (2), 127-130 (2001)
  MEDLINE
    21475601
REFERENCE
  2 (bases 1 to 1529)
  AUTHORS
    Suzuki,M., Morimatsu,M. and Syuto,B.
  TITLE
    Direct Submission
  JOURNAL
    Submitted (23-NOV-2000) Masako Suzuki, Iwate University, Department
    of Veterinary Medicine, Faculty of Agriculture, Ueda 3-18-8,
    Morioka, Iwate 020-8550, Japan (E-mail:u9998001@iwate-u.ac.jp,
    Tel:81-19-621-6212)
  LOCATION/QUALIFIERS
    1..1529
    /organism="Bos taurus"
    /db_xref="taxon:9913"
    /tissue_type="liver"
  1..1529
    /gene="cbp b04"
    13..1431
    /gene="cbp b04"
    /codon_start=1
    /product="chitin binding protein b04"
    /protein_id="BAB71805.1"
    /db_xref="GI:16754895"
    /translation="MAKLIFLFLAFLNAQLGSAQLVCFNSNAQYRPGSGKPKPD
    NIDPCLHLIYAFAGMSNEITTEIWNDAVLSYFNLDLKKNSQLKILATGGWNFG
    TAPFAMVATPENRKTFSVVKFLHOYDGLDFDEWEPGPRGSPQDKHLFTVLVQ
    ETRFAEQEAKTKPRLLVTAAGAAGLSNIOAGYEIPOLSOYLDIFHWMTYDFHGSW
    EGYTGENSEPLKYPDTGNTLVNEYAMNYKKNKAPAEKILIGFPAYGHNFILRDA
    SNNGIGATSGAGPAPYTRAGFWAYEICAFKDKGDGATEAWDDSONPYATKGTGW
    GYDYNYSFRKAKWLKNFNGGAMWAILDDFTGTCNOGKPEPLINTLKLKLSA
    TCNASTQSPNSPNSGSGKSSSSSEGRGCGACADGLYPVADNRNRAFNCVNGI
    TVKONCLTGLVFDTSCHCNMA"
    1508..1513
    polyA_signal
  -BASE COUNT
    391 a 405 c 365 g 368 t
  ORIGIN
    Query Match 62.9%; Score 959.2; DB 4; Length 1529;
    Best Local Similarity 80.0%; Pred. No.1.5e-287;
    Matches 1142; Conservative 0; Mismatches 283; Indels 3; Gaps 1;
    QY 1 ATGGCAAGCTACTTCTCTGCTCAGAGCTCTGGCTCTTCTGCTGAATGCTCAGCTGGGGTCT 60
```


186 GAACAATGAGATCACCAACATAGAAATGGAATGATGTTACTCTCTATGAAGCTTTCAATGA 245
18 GCACGAACAGGTGCGCGACTCTGGAGCCAGCGTGTTCAGTCTGGTGAATC 77
246 CTTGAAAACAGGACAGCAACAACTGAAACCCCTCTGGCAATGAGGCTGGAATTTGG 305
78 CTCCATAGCTGGAAGAGCCAGCTGAAACCTCTCTGGCCATTTGAGGCTGGAATTTGAG 137
306 AACTGCTCTTCTACTACCATGGTTTCCACTTCTCAGAACCCGAGAGCTTTCATACCTC 365
138 GACTGCCCTTCTACTGCCATGGTTTCTACTCTTGAGAACCCGAGAGCTTTCATACCTC 197
366 AGTCATCAATTTCTGGCTCAGTATGGGTTTGTATGGACTGGACCTGGAGTATGGAATACCC 425
198 AGTCATCAATTTCTGGCCAGTATGAGTTTGGAGGCTGGACTTGGAGTATGGAATACCC 257
426 AGGCTCAGCTGGGAGCCCTCTCAGGACAAACATCTCTTCACTGCTCTGGTGAAGGAAT 485
258 TGGCTCTCGTGGGAGCCCTCTCAGGACAAACATCTCTTCACTGCTCTGGTGAAGGAAT 317
486 GCGTGAAGCTTTTGGAGGAGGCTATTTGAGCAACAGGCCAGAGCTGATGTTTACATG 545
318 GCGTGAAGCTTTTGGAGGAGGCTCAGCAGATCAACAGGCCAGGCTGATGTTTACATG 377
546 TGCTGTAGCTGGTGGGATTTCCAAATCCAGGCTGGCTATGAGATCCCTGGAATTTCTAA 605
378 TGCAGTAGCTGCTGCACTCTCAATATCCAGTCTGGCTATGAGATCCCTGGAATTTCTAA 437
606 GTACTGTGATTTCAATCATGATCATGATGATGATGATGATGATGATGATGATGATGATG 665
438 GTACTGTGATTTCAATCATGATCATGATGATGATGATGATGATGATGATGATGATGATG 497
666 TGGGAGGAATAGTCTCTTTTCAAAATACCTTACTGAGACTGATGATGATGATGATGATG 725
498 TGGAGAGACAGCCCTCTCAAAATACCCAGTATGATGATGATGATGATGATGATGATGAT 557
726 TGTGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 785
558 TGTGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 617
786 TGGATTTCCAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 845
618 TGGATTTCCAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 677
846 TGCCCTTCACTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 905
678 TGCCCTTCACTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 737
906 CTACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 965
738 TTACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 797
966 AGAAGTCCCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1025
798 GGAAGTCCCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 857
1026 CAGTGTAAAGCTCAGTGGCTTAAAGCAACAAATTTGGAGGTGATGATGATGATGATGATG 1085
858 CGATATTAAGCTCAATGGCTTAAAGCAACAAATTTGGAGGTGATGATGATGATGATGATG 917
1086 TGACCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1145
918 TGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 977
1146 TTTGACAAACCCCTTGGCATATGATGATGATGATGATGATGATGATGATGATGATGATG 1205
978 CCTGAAGAGAGCCCTTGGCCCTGAGAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGC 1037
1206 GCCAGTACTACTCTCTC-----AGGAAGTGGAGTGGGTTGAACTCCGGAGG 1256
1038 GCCAATACTCTCTCTCCAGTGGCAGCGGGAACGGGAGCGGGAGTAGAGCTCTGGAGG 1097

QY 1257 AAGCTCTGGAGCAGTGGATTTCTGTGCGGACAAACAGAGATGGCTCTTACCCTGTGGCAGA 1316
Db 1098 CAGCTCGGAGGAGTGGATTTCTGTGCGGACAAACAGAGATGGCTCTTACCCTGTGGCAGA 1157
QY 1317 TGACAGAAATGCTTTTGGCAGTGCATCAATGGAATACATACAGCAGCATTTGTCAAGC 1376
Db 1158 TAACAGAAATGCTTTTGGCAGTGCATCAATGGAATACATACAGCAGCATTTGTCAAGC 1217
QY 1377 AGGCTCTTTTGTATACAGCTGTAATTTGTGCAACTGGCAGTGAACCTAA 1428
Db 1218 CGGCTGTCTTCGACACAGCTGTGATGTTGTCGACATGGGATAAACCTGA 1269

RESULT 8
AX405989
LOCUS AX405989 1369 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 404 from Patent WO0222660.
ACCESSION AX405989
VERSION AX405989.1 GI:21439417
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1
AUTHORS Tang, Y. T., Liu, C., Zhou, P., Asundi, V., Zhang, J., Zhao, Q. A., Ren, F.,
Xue, A. J., Yang, Y., Wehrman, T. and Drmanac, R. T.
TITLE Novel nucleic acids and polypeptides
JOURNAL Patent: WO 0222660-A 404 21-MAR-2002;
HYSEQ, INC. (US)
FEATURES
Location/Qualifiers
1..1369
/organism="Homo sapiens"
/db_xref="taxon:9606"
154..1260
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD35096.1"
/translation="MVSPTENROTFTISVIFKLOYEFDGLDFDWEYPSRGSPPQDK
HLFTVLQEMREAEQEKAKOINKPRLAWTAVAAGISNIQSGYEIPQLSQYLDVTHVM
TYDLHGSWEYTGNSPLIKYPTDTSNAYLVNDVYMNWKDNGAPAKLIVGPEPTYG
HNFILSNPSTGIGATGSGAGPAGPYAKESGWAYEICTELKNGATGQWAPQEVY
AYQNVWVYDNIKSFDAQKHLKHFNGAMVAIDLDDFTGFCNGKRPFLSTLK
KALGLQASCTAPAOPTPIPTAAPSNGSGSSSSSGSGSGSFCAVRANGLYPVAN
NRNFAHCVNGVTYQNCQAGLVFDTSCDCNWA"
BASE COUNT 327 a 393 c 341 g 308 t
ORIGIN

Query Match 56.0%; Score 854; DB 6; Length 1369;
Best Local Similarity 82.6%; Pred. No. 1.1e-254;
Matches 993; Conservative 0; Mismatches 200; Indels 9; Gaps 1;

QY 236 CTTTCAATGACTTGAAAACAGGACAGCAACAACTGAAAACCCCTCTCGCAATTTGAGGCT 295
Db 65 CTTTACTGAAATTTGACGCTCGGACACGCCAGCTGAAAACCTCTCTCGCCATTTGAGGCT 124
QY 296 GGAATTTGGAACTGCTCTTTTCACTACCATGTTTCCACTTCTCAGAACCCGAGACCT 355
Db 125 GGAATTTGGAACTGCTCTTTTCACTACCATGTTTCCACTTCTCAGAACCCGAGACCT 184
QY 356 TCATACCTCAGTCAATCAAAATTTTGCCTCAGTATGGTTTGTGATGGACTGGACCTGGACT 415
Db 185 TCATACCTCAGTCAATCAAAATTTTGCCTCAGTATGGTTTGTGATGGACTGGACTGGACT 244
QY 416 GGAATACCCAGGCTCAGTGGAGCCCTCTCAGGACAAAGATCTCTTCACTGTCCTGG 475
Db 245 GGAATACCCAGGCTCAGTGGAGCCCTCTCAGGACAAAGATCTCTTCACTGTCCTGG 304
QY 476 TGAGGAATACCCAGGCTTGTGAGCAGGAGGCTTATTTGAGAGCAACAGGCCCCACACTGA 535
Db 305 TGCAGGAATACCCAGGCTTGTGAGCAGGAGGCTTATTTGAGAGCAACAGGCCCCACACTGA 364

QY	536	TGTTACTGCTGCTAGCTGGTGGGATTTCCAAATCCACAGCTGGGTGCTATGAGATCCCTG	595
Db	365	TGTTACTGCTGCTAGCTGGTGGGATTTCCAAATCCACAGCTGGGTGCTATGAGATCCCTG	424
QY	596	AACTTCTAAGTACCTGGATTTTCATCCATGTCATGACATATGACCTCATGGCTCTCTGGG	655
Db	425	AACTGTACAGTACCTGGACATACATCCATGTCATGACCTAGACCTCATGGCTCTCTGGG	484
QY	656	AGGGCTACACCTGGGAGAAATAGTCTCTTTTACAATAACCTTACCTGAGACTGGTAGCAATG	715
Db	485	AGGGCTACACTGGAGAGAACAGCCCTCTTACAATAACCTGAGACTGGTAGCAATG	544
QY	716	CCTACTCAATGTGATTTGATGAACTATTTGGAAGAACAAATGGAGCCGCCAGCTGAGA	775
Db	545	CCTACTCAATGTGATTTGATGAACTATTTGGAAGAACAAATGGAGCCGCCAGCTGAGA	604
QY	776	AGCTCATTTGGATTTCCAGAGTAGGACACACCTTCATCTCTGAGAAACCCCTCTGTATA	835
Db	605	AGCTCATTTGGATTTCCAGAGTAGGACACACCTTCATCTCTGAGAAACCCCTCTGTATA	664
QY	836	ATGGAAATTTGGTCCCTTACCTCTGTGTATGCTGCTGGCCCTATACAGACAGGCTG	895
Db	665	CTGGAAATTTGGTCCCTTACCTCTGTGTATGCTGCTGGCCCTATACAGAGGCTG	724
QY	896	GGTTCTGGCCCTACTATGAGATTTTGCACCTTTCTGAGAAGTGGAGCCACCTGAGTCTGGG	955
Db	725	CGATCTGGCTTACTACGAGATCTGTACCTTCTTGAAGAAATGGAGCCACCTGAGGATGG	784
QY	956	ATGCTCTCCCAAGAGTGCCTTATGCTATATAGGCCACAGAGTGGCTTGGCTATGACATTA	1015
Db	785	ATGCTCTCCCAAGAGTGCCTTATGCTATATAGGCCACAGAGTGGCTTGGCTATGACATTA	844
QY	1016	TCAAGAGCTTTCAGTCTTAAGGCTCAGTGGCTTAAAGCAGAGAACAAATTTTGGAGTGGCATGA	1075
Db	845	TCAAGAGCTTTCAGTCTTAAGGCTCAGTGGCTTAAAGCAGAGAACAAATTTTGGAGTGGCATGA	904
QY	1076	TTCTGGCCATGACCTTGATGACTTCATCTGCTCTTCTTGTGATCAGGGAAAAATTTCCCTC	1135
Db	905	TTCTGGCCATGATCTGGATGACTTCATCTGCTCTTCTTGTGATCAGGGAAAAATTTCCCTC	964
QY	1136	TGACTTCTACTTTGAAACAAAGCCCTTGGCATATCCATGAAGTGTGACAGCTCCTTGAGC	1195
Db	965	TAACTTCCACCTTGAAGAAAGCCCTTGGCATATCCATGAAGTGTGACAGCTCCTTGAGC	1024
QY	1196	TGCTTCTCGGACGAGTACTACTCTCTCC-----AGGAAGTGGGAGTGGGGTGGAA	1246
Db	1025	AGCCATTTAGCCCAATACTCTCTCCAGTGGCAGCGGAGACGGAGCGGAGTAGCA	1084
QY	1247	GCTCGGAGGAGCTCTGGAGGAGTGGATCTGTGCGGCAAGACAGATGGCTCTTACC	1306
Db	1085	GCTCTGGAGGAGCTCTGGAGGAGTGGATCTGTGCTGTGACAGCAAGCGGCTCTACC	1144
QY	1307	CTGTGGCAGATGACAGAAATGCTTTTGGCAGTGCATCATGAAATCAGATCCAGCAGC	1366
Db	1145	CGGTGGCAAAATACAGAAATGCTTCTGGCACTGCGTGAATGGAGTACAGTACAGCAGA	1204
QY	1367	ATTGTCAAGCAGGCTCTTTTTCATACCACTGTAATTCGTCACTGGCCCACTGAACCT	1426
Db	1205	ACTGCCAGCGGCTGTCTTCACACAGGCTGTGATGCTGCACTGGCCCACTGAACCT	1264
QY	1427	AA 1428	
Db	1265	GA 1266	
RESULT 9			
M94584			
LOCUS			
DEFINITION			
M94584 Mus musculus secretory protein precursor (Yml) mRNA, complete cds.			
ACCESSION			
M94584.2			
VERSION			
M94584.2 GI:11140876			
KEYWORDS			
Mus musculus.			
SOURCE			
M94584			
LOCUS			
DEFINITION			
M94584 Mus musculus secretory protein precursor (Yml) mRNA, complete cds.			
ACCESSION			
M94584.2			
VERSION			
M94584.2 GI:11140876			
KEYWORDS			
Mus musculus.			
SOURCE			
M94584			
LOCUS			
DEFINITION			
M94584 Mus musculus secretory protein precursor (Yml) mRNA, complete cds.			
ACCESSION			
M94584.2			
VERSION			
M94584.2 GI:11140876			
KEYWORDS			
Mus musculus.			
SOURCE			
M94584			
LOCUS			
DEFINITION			
M94584 Mus musculus secretory protein precursor (Yml) mRNA, complete cds.			
ACCESSION			
M94584.2			
VERSION			
M94584.2 GI:11140876			
KEYWORDS			

```
Db 1153 TCTACTTTAAAGGAGATCTCAATATACACAGTGCAGTTGCAAGGCGCTTA 1205
RESULT 10
D87757
LOCUS D87757 1506 bp mRNA linear ROD 12-JAN-2000
DEFINITION Mus musculus mRNA for ECF-L precursor, complete cds.
ACCESSION D87757
VERSION D87757.2 GI:6015435
KEYWORDS ECF-L precursor.
SOURCE Mus musculus bone marrow cDNA to mRNA.
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 1506)
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE Identification of a novel eosinophil chemotactic cytokine (ECF-L)
JOURNAL 1 (bases 1 to 1506)
MEDLINE Ohashi M, Arita H, and Hayai N.
REFERENCE 2 (bases 1 to 1506)
AUTHORS J. Biol. Chem. 275 (2), 1279-1286 (2000)
TITLE Identification of a novel eosinophil chemotactic cytokine (ECF-L)
JOURNAL 2009202
AUTHORS Ohashi M.
TITLE Direct Submission
JOURNAL Submitted (06-SEP-1996) Makoto Ohashi, The University of
Tokushima, Faculty of Integrated Arts and Sciences; 1-1
Minami-Johsanjima, Tokushima, Tokushima 770, Japan
(E-mail: ohashi@ias.tokushima-u.ac.jp, Tel: 0886-56-7261,
Fax: 0886-56-7298)
COMMENT On Oct 7, 1999 this sequence version replaced gi:1545818.
Sequence updated (05-Oct-1999).
FEATURES
source
1. 1506
/organism="Mus musculus"
/db_xref="taxon:10090"
/tissue_type="bone marrow"
gene
1. 1506
/organism="Mus musculus"
/db_xref="taxon:10090"
CDS
1. 1197
/organism="Mus musculus"
/db_xref="taxon:10090"
/translation="MAKILVLTGLAILNVQLSSQYLMYYTWSAKDRPIEGSEKPG
NIDPCLTHTLIVAPQMONNETTYTHEODLRYEALNGKDKNTKELTKLLATIGGKFG
PASFSAMYSTPQNFIOFISVTRFLRNFEDLNLDWOYPGSRGPPDKHLFSLVK
EMRAFEESVEKDIPLLSTGAGIDIVKSGYKIFELSSLDYIQWYDLHDPK
DGYTGNSPLYSKYDYGKADNLNVDLSIISYKWDHGAASEKLIVFPAYGHTFLLSDP
SKTGIGAPTISTGPGKYDEGLAYEVCFLNEGATEVMDAPQEPYAYQGNHW
GYDNRVSFKLKAQWLKDNLGVVWPLDMDFSGSFCHQHFPLTSLKGLDNLHSA
SCKGPY"
BASE COUNT 423 a 338 c 333 g 412 t
ORIGIN
Query Match 46.8%; Score 713; DB 10; Length 1506;
Best Local Similarity 74.9%; Pred. No. 1.2e-210;
Matches 893; Conservative 0; Mismatches 300; Indels 0; Gaps 0;
QY 1 ATGGCCAGCTACTTCTCGTCACAGGTCTGGCTCTTCTGCTGAATGCTCAGCTGGGTCT 60
Db 1 ATGGCCAGCTACTTCTCGTCACAGGTCTGGCAATTTCTTCTGAAGCTACAGCTGGGTCT 60
QY 61 GCCTACATCATGATGCTATTTACCAACTGGGCCCCCATCTCGCCAGGTCTGGGAGC 120
Db 61 TCCTACAGCTGATGCTGCTACTATACAGTTGGGTAAAGGACAGGCCAATAGAGGAGT 120
QY 121 TTCAAGCTGATGACATTAACCCCTGCTGCTACTCAGCTCATCTATGCTCTTCTGCTG 180
Db 121 TTCAAGCTGATGATTAATGACCCCTGCTGCTACTCAGCTCATCTATGCTCTTCTGCTG 180
QY 181 ATGCAGAAATAGATGATCACCACCATAGAAATGATGTTTCTCTATAAAGCTTTC 240
Db 181 ATGCAGAAATAGATGATCATTACACATGACGAAGACTTTCGCTGACTTATGAAGCATTTG 240
```

```
Db 73 TCCTACAGCTGATGCTGCTACTATACAGTTGGGCTAAGGACAGGCCAATAGAGGAGT 132
QY 121 TTCAGGCTGATGACATTAACCCCTGCTGCTGCTACTCAGCTGATGATGCTCTGCTGG 180
Db 133 TTCAAGCTGATGATTAATGACCCCTGCTGCTGCTACTCAGCTGATGATGCTCTGCTGA 192
QY 181 ATGCAGAAATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 240
Db 193 ATGCAGAAATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 252
QY 241 AATGACTTGAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
Db 253 AATGCTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 312
QY 301 TTGGAAGCTGCTCTTCACTACCATGCTTCCACTTCTCAGAACCGGACGACCTTCATT 360
Db 313 TTGGAAGCTGCTCTTCACTACCATGCTTCCACTTCTCAGAACCGGACGACCTTCATT 372
QY 361 ACCTCAGTCATCAAAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
Db 373 CAGTCAGTTATCAGATTCCTTCTGCTCAATATATATATATATATATATATATATATAT 432
QY 421 TACCAGGCTCAGTGGAGGCTCTCTCAGGACAGAGATCTCTTCACTGCTCTGCTGCTG 480
Db 433 TACCCTGGTCTCGAGGAGGCTCTCTTCAAGGACAAACATCTCTTCACTGCTCTGCTG 492
QY 481 GAAATGCTGAGGCTTTGAGCAGGAGCTATTGAGACCAAGCCGACGACTGATGTT 540
Db 493 GAAATGCTGAGGCTTTGAGGAGGAGCTTTGAGGAGGAGGAGGAGGAGGAGGAGGAG 552
QY 541 ACTGCTGCTGAGTGGTGGGATTTCCAACTCCAGGCTGGCTATGAGATCCCTCAACTT 600
Db 553 ACTTCCAGAGGAGGAGGATCATTGACGTATCAAGTCTGGGTACAGAGATCCCTGACTG 612
QY 601 TCTAAGTACCTGAGATTCATGATGATGATGATGATGATGATGATGATGATGATGATG 660
Db 613 TCTCAGTCTCTTACATATATTCAGGCTATGATGATGATGATGATGATGATGATGATG 672
QY 661 TACACTGGGAGATGCTCTTTTACAAATACCTCTGCTGAGCTGGTAGCAATGCTTAC 720
Db 673 TACACTGGGAGATGCTCTTTTACAAATACCTCTGCTGAGCTGGTAGCAATGCTTAC 732
QY 721 CTCAATGGGATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
Db 733 CTCAATGGGATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 792
QY 781 ATGTTGGATTTCCAGAGATGAGGACACCTTCCTGAGAAACCCCTCTGATATGGA 840
Db 793 ATGTTGGATTTCCAGAGATGAGGACACCTTCCTGAGAAACCCCTCTGATATGGA 852
QY 841 ATGTTGGATTTCCAGAGATGAGGACACCTTCCTGAGAAACCCCTCTGATATGGA 900
Db 853 ATGTTGGATTTCCAGAGATGAGGACACCTTCCTGAGAAACCCCTCTGATATGGA 912
QY 901 TGGGCTCTACTGATGATTTGACCTTCTGAGAGAGGAGGAGGAGGAGGAGGAGGAGG 960
Db 913 CTGGCTCTACTGATGATTTGACATTTCTGATGATGATGATGATGATGATGATGATG 972
QY 961 TCCCAAGAGTCCCTATGCTCTTAAAGGCAACAGAGTGGCTTGGCTATGACAAATACAG 1020
Db 973 CCCCAGGAAGTACCTATGCTCTTAAAGGCAACAGAGTGGCTTGGCTATGACAAATACAG 1032
QY 1021 AGCTTCAGTGTAAAGGCTTAAAGGCTTAAAGGCAACAGAGTGGCTTGGCTATGACAA 1080
Db 1033 AGCTTCAGTGTAAAGGCTTAAAGGCTTAAAGGCAACAGAGTGGCTTGGCTATGACAA 1092
QY 1081 GCCATTGACCTTGAAGCTTCACTGCTCTTCTGATGATGATGATGATGATGATGATG 1140
Db 1093 CCCCTGGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1152
QY 1141 TCTACTTGAAGAAAGGCTTGGGATATCCACTGAGGTTGACAGGCTCTCTGA 1193
```


Db 241 AATGGTCTGAAGACAGGACACACAGCTAGAGTAAAACTCTCTCGCCATGGAGGATGGAAG 300
Qy TTTGGAACCTGCTCTTTTACACACACATGCTTCCACTTCTCAGAACCGGACACACTTCATT 360
Db TTTGGAACCTGCTCTTTTACACACACATGCTTCCACTTCTCAGAACCGGACACACTTCATT 360
Qy ACCCTGACATCAAAATTTCTGCTGATGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 420
Db AAGTCAGTATACAGATCTCTGCTCAATATCAATTTGATGGCTCAACCTGACCTGGCAG 420
Qy TACCCAGGCTCAGCTGGGAGGCTCTCTCAGACAGACATCTCTTCACTGCTCTGCTGGTGAAG 480
Db TACCTGGGCTCAGGAGGCTCTCTTAAAGACAAACATCTCTTCACTGCTCTGCTGGTGAAG 480
Qy GAAATGCTGTAAGCTTTTGAAGGAGGCTTATGAGAGCAACAGGCCAGACTGATGTT 540
Db GAAATGCTGTAAGCTTTTGAAGGAGGCTTATGAGAGCAACAGGCCAGACTGATGTT 540
Qy ACTGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 600
Db ACTTCCAGGAGCTGGATTCATTCAGCTAATCAAGTCTGGGTACAAGATCCCTGAAGTGGC 600
Qy TCTAAGTACCTGGATTCATTCAGCTAATCAAGTCTGGGTACAAGATCCCTGAAGTGGC 660
Db TCTAAGTACCTGGATTCATTCAGCTAATCAAGTCTGGGTACAAGATCCCTGAAGTGGC 660
Qy TACACTGGGAGATAGCTCTTTTACAAATACCTCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 720
Db TACACTGGGAGATAGCTCTTTTACAAATACCTCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 720
Qy CTCAATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 780
Db CTCAATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 780
Qy ATTGTTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 840
Db ATTGTTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 840
Qy ATGTTGCTGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Db ATGTTGCTGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Qy TGGGCTTACTATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
Db CTGGCTTACTATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
Qy TCCCAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
Db ACCCAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
Qy AGCTTCACTGTTAAGGCTCAGTGGCTTAAAGCAACAAATTTTGAAGGCTGCTGCTGCTGCTGCT 1080
Db AGCTTCACTGTTAAGGCTCAGTGGCTTAAAGCAACAAATTTTGAAGGCTGCTGCTGCTGCTGCT 1080
Qy GCCATGACCTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
Db CCCCTGGACATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
Qy TCTACTTTGAACAAAGCCCTTGGCATATCCACTGAGGTTGCAAGCTGCTGCTGCTGCTGCTGCT 1193
Db ACTACTTTAAGAGAGATCTGAATGTACACAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGC 1193

RESULT 12
AY065557 LOCUS AY065557 1209 bp mRNA linear ROD 22-MAY-2002
DEFINITION Mus musculus putative secretory protein precursor (Ym2) mRNA,
complete cds.
ACCESSION AY065557
VERSION AY065557.1 GI:18086513
KEYWORDS Mus musculus.
SOURCE Mus musculus.
ORGANISM Mus musculus.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1209)
Chang,N.-C.A.
Tissue-specific Expression of Ym2 Protein
Unpublished
2 (bases 1 to 1209)
Chang,N.-C.A.
Direct Submission
Submitted (05-DEC-2001) Institute of Microbiology & Immunology,
National Yang-Ming University, #155 Sec. 2 Li-Nong St., Beitou,
Taipei 112, Taiwan
Location/Qualifiers
1. 1209
/organism="Mus musculus"
/strain="BALB/c"
/db_xref="taxon:10090"
/clone="pGEMT-ez-ym2.s"
/tissue_type="stomach"
1. 1209
/gene="Ym2"
1. 1209
/gene="Ym2"
/codon_start=1
/product="putative secretory protein precursor"
/protein_id="AAL57751.1"
/db_xref="GI:18086514"
/translation="MAKILIVTGLAILNVOLGSSYOLMCVYTSWAKDRPTEGSFKPG
NIDPLCLTHLIYAFAGKNNETILSEQDLNEDYALNGLKDRNETLTKLLAIGWKKFG
PAFSSMYSTPQNRQTFIKSVIRFLQYNFNLDMQYDGPGRSPGSLKHLFSLVQ
EMKAFEESELYNHPRLLSSTGAGFIDVYKSGYKIPELSQSLDYIQVYHTFLHDPK
NGYTGENSELYKSPYDYGKADLVDSIITWKDHGAASEKLIYGFAYGHTFLLSDP
SKNGIGDPTVSAGPPGKYTNEGGLAYFEICTELNEGATEIFDQTEVPYAVLGNWV
GIDNVRSEFKLKAQWLKDNGLGAVVWPLDMDDFSGSFCHQGRFPLTTLTKRDLNVHSA
CYKASIRGEL"
326 a 274 c 282 g 327 t
BASE COUNT
ORIGIN
Query Match 46.3%; Score 706.6; DB 10; Length 1209;
Best Local Similarity 74.5%; Pred. No. 1.1e-208;
Matches 889; Conservative 0; Mismatches 304; Indels 0; Gaps 0;
Qy 1 ATGGCCAAAGCTACTTCTCGTCACAGGCTGCGCTCTTCTGCTGAATGCTCAGCTGGGCTCT 60
Db 1 ATGGCCAAAGCTACTTCTTGTTCACAGGCTGCTGGCAATCTTCTGAATGTACAGCTGGGCTCT 60
Qy 61 GCCTACATATCATATGCTATTTACCAACTGGGGCCAGATCGCCAGGCTCGGGAGC 120
Db 61 TCCTACAGCTGATGCTGCTACTATACAGCTGGGCTAAGACAGCCCAAGAGGAGT 120
Qy 121 TTCAGCCCTGATGACATTAACCCCTGCTGCTACTACCTGATCTATGCTGCTGG 180
Db 121 TTCAGCCCTGATGATTAATTTGACCCCTGCTGCTACTACCTGATCTATGCTGCTGG 180
Qy 181 ATGCAGAAATATGATGATCACCACCAATAGAAATGATGTTACTCTCTATAAGCTTTC 240
Db 181 ATGAAGAATATGATGATCATTACTTAAATGAGCAAGAGCTTCGCTGACTATGAAGCATTA 240
Qy 241 AATGACTTGAACAAACAGCAACAGCAACCTTCCGGCAATTTGGAGCTTGGAGC 300
Db 241 AATGCTTGAACAAACAGCAACAGCAACCTTCCGGCAATTTGGAGCTTGGAGC 300
Qy 301 TTTGGAACCTGCTCTTTTCACTACCATGTTTCCACTTCTCAGAACCGGACACACTTCATT 360
Db 301 TTTGGAACCTGCTCTTTTCACTACCATGTTTCCACTTCTCAGAACCGGACACACTTCATT 360
Qy 361 ACCTCAGTCATCAAAATTTCTGCTGCTGATGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 420
Db 361 AAGTCAGTATACAGATCTCTTCTGCTCAATATCAATTTGATGGCTCAACCTGACCTGGCAG 420
Qy 421 TACCCAGGCTCAGCTGGGAGGCTCTCTCAGGACAGACATCTCTTCACTGCTCTGCTGGTGAAG 480
Db 421 TACCCAGGCTCAGGAGGCTCTCTTAAAGCAACAACTCTCTTCACTGCTCTGCTGGTGAAG 480

481 GAAATGCGTGAAGCTTTTGAGCAGGAGGCTATTGAGAGCAACAGGCCAGACTGATGGTT 540
|||||
481 GAAATGCGTGAAGCTTTTGAGGAGGAATCCACTTTGAACACACATTCGAAGCTGCTACTC 540
|||||
541 ACTGCTGCTGCTAGCTGGTGGATTTCCACATCCAGCTGGCTGATGAGATCCCTGAACTT 600
|||||
541 ACTTCCACAGAGCTGGATTCATTTGACGTAATCAAGTCTGGGTACAGATCCCTGAACTG 600
|||||
601 TCTAAGTACTGATTCATCCATGTCATGACATATGACCTCCATCCAGCTGCTGGAGGCG 660
|||||
601 TCTAAGTACTGATTCATCCATGTCATGACATATGACCTCCATCCAGCTGCTGGAGGCG 660
|||||
661 TACACTGGGAGAAATAGTCTCTTTTACAAATACCCTACTGAGACTGGTAGCAATAGCTAC 720
|||||
661 TACACTGGGAGAAATAGTCTCTCTTATAATCTCCATATGACATTTGGAAGAGTCTGAT 720
|||||
721 CTCATGTTGGATTTATGTCATGACTATTGGAAGAACAAATGGAGCCCGCTGAGAGCTC 780
|||||
721 CTCATGTTGGATTTATGTCATGACTATTGGAAGAACAAATGGAGCCCGCTGAGAGCTC 780
|||||
781 ATTGTTGGATTTCCAGAGTATGACACACACCTTCTGAGAAACCCCTCTGATAATGGA 840
|||||
781 ATTGTTGGATTTCCAGAGTATGACACACACCTTCTGAGTAACTCTTATGAGTAACTGGA 840
|||||
841 ATTGTTGGCTCCCTACCTCTGCTGATGAGGCTGCTGGCCCTATACAGAGAGGCTGGTTC 900
|||||
841 ATAGGTGACCTACTGTTAGTCTGGAGCACCAGGAAGTACACAAATGAACAGGACTC 900
|||||
901 TGGGCTCTACTATGATTTGACACTTTCTGAGAAAGTGGAGCCACTGAGGTCTGGATGCC 960
|||||
901 CTGGCTTACTTTGAGATTTGATGATTTCTGATGAAAGGAGCCACTGAGATCTTTGATGCC 960
|||||
961 TCCAAAGAGTGGCTCTGCTATGAGGCAACAGGCTGGCTTGGCTGATGACAAATATCAAG 1020
|||||
961 ACCAGGAGTACCTGATGCTATCTGGGTAATGAGTGGTGGTATGACATGTCAGG 1020
|||||
1021 AGCTTCAAGTGAAGGCTCAGTGGCTCAAGGACAAATTTAGGAGTGGCTGGTCTGG 1080
|||||
1021 AGCTTCAAGTGAAGGCTCAGTGGCTCAAGGACAAATTTAGGAGTGGCTGGTCTGG 1080
|||||
1081 GCCATTGACCTGATGATCTGCTGCTGCTGCTGCTGATCAGGAGAAATTTCTCTGACT 1140
|||||
1081 CCCTGACATGATGATCTGCTGCTGCTGCTGCTGCTGATCAGGAGAAATTTCTCTGACT 1140
|||||
1141 TCTACTTTGAACAAAGCCCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1193
|||||
1141 ACTACTTTAAAGAGAGATCTGAATGTACACAGTGCAAGTTGCAAGGCCCTCTTA 1193
|||||

RESULT 13

AK098814
LOCUS
DEFINITION
Homo sapiens cDNA FLJ25948 fis, clone STW04207, highly similar to
Homo sapiens acidic mammalian chitinase precursor.
AK098814
ACCESSION
VERSION
KEYWORDS
SOURCE
oligo capping; fis (full insert sequence).
Homo sapiens stomach mucosa cDNA to mRNA, clone_lib:STM
clone:STM04207.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1
Ninomiya, K., Nagatsuma, M., Kanda, K., Kondo, H., Yokoi, T.,
Kodaira, H., Furuya, T., Takahashi, M., Kikkawa, E., Omura, Y., Abe, K.,
Kamihara, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M.,
Suzuki, Y., Hata, H., Nakagawa, K., Mizuno, S., Morinaga, M.,
Kawanura, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H.,
Nishikawa, T., Sugiyama, A., Kawakami, B., Nagai, K., Isogai, T. and
Sugano, S.

NEDO human cDNA sequencing project

Unpublished

TITLE

JOURNAL

REFERENCE

AUTHORS
TITLE
JOURNAL

COMMENT

FEATURES

source

BASE COUNT

ORIGIN

Query Match 43.9%; Score 669; DB 9; Length 4250;
Best Local Similarity 81.2%; Pred. No. 7.8e-197;
Matches 793; Conservative 0; Mismatches 175; Indels 9; Gaps 1;

2 (bases 1 to 4250)
Sugano, S. and Suzuki, Y.
Direct Submission
Submitted (08-JUL-2002) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Laboratory of Genome Structure, Human Genome
Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
(E-mail: cdna@ims.u-tokyo.ac.jp, Tel: 81-3-5449-5286,
Fax: 81-3-5449-5416)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction and 5'-end one pass sequencing: Institute of Medical
Science, University of Tokyo, Laboratory of Genome Structure, Human
Genome Center; 3'-end one pass sequencing: RAB; clone selection for
full insert sequencing: RAB and Helix Research Institute.

Location/Qualifiers

1..4250

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="STM04207"

/tissue_type="stomach mucosa"

/clone_lib="STM"

/note="cloning vector: pME18SFL3"

BASE COUNT 1262 a 868 c 966 g 1154 t

461 TCTTCTACTGCTGCTGGAAGAAATCGGTGAAGCTTTTGAGCAGGAGGCTATTGAGAGCA 520
|||||

3184 TATGCCCTTATTATCTGTAGGAATCGTGAAGCTTTTGAGCAGGAGGCTAAGCAGATCA 3243
|||||

521 ACAGGCCAGACTGATGTTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 580
|||||

3244 ACAAGCCAGGCTGATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3303
|||||

581 GCTATGAGATCCCTGAACTTTCTTAAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 640
|||||

3304 GCTATGAGATCCCTGAACTTTCTTAAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3363
|||||

641 TCCATGGCTCCCTGGGAGGCTACACTGGGAGATAGTCTCTCTTACAAATACCCCTACTG 700
|||||

3364 TCCATGGCTCCCTGGGAGGCTACACTGGGAGATAGTCTCTCTTACAAATACCCCTACTG 3423
|||||

701 AGACTGGTAGCAATGCTTACCTCAATGTGGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 760
|||||

3424 ACACCGCAGCAAGCTTACCTCAATGTGGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3483
|||||

761 GAGCCCCAGCTGAGAGCTCAATTTGGATTTCCAGAGTATGGACACACCTTCATCCTGA 820
|||||

3484 GAGCACCAGCTGAGAGCTCAATTTGGATTTCCAGAGTATGGACACACCTTCATCCTGA 3543
|||||

821 GAAACCCCTCTGATTAATGGAATTTGGTGGCCCTTACCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 880
|||||

3544 GCAACCCCTCTGATTAATGGAATTTGGTGGCCCTTACCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 3603
|||||

881 ATACCCAGCAGCTGGGTTCTGGCCCTTACTATGAGATTTGCACCTTTCTGAGAAAGTGGAG 940
|||||

3604 ATGCCAAGAGCTGGGATCTGGGCTTACTACGAGATCTGTACCTTCTGAGAAAGTGGAG 3663
|||||

941 CCAGTGGTCTGGGATGCTCCCAAGAGTGGCCCTTATGCTTATAGGCCAACAGCAGTGGC 1000
|||||

3664 CCAGTGGTCTGGGATGCTCCCAAGAGTGGCCCTTATGCTTATAGGCCAACAGCAGTGGC 3723
|||||

1001 TTGGCTTATGACATATCAAGAGCTTCAGTGTGTTAAGCTCAGTGGCTTAAGCAGCAACAAT 1060
|||||

3724 TTGGCTTATGACATATCAAGAGCTTCAGTGTGTTAAGCTCAGTGGCTTAAGCAGCAACAAT 3783
|||||

1061 TTGGAGGTGCCATGATCTGGGCCATGACCTTGTGATCTTCACTGCTCTTTCTGTGATC 1120
|||||

3784 TTGGAGGTGCCATGATCTGGGCCATGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTG 3843
|||||

QY	1121	AGGGAATTTCTCTGACTTCTACTTTGAACAAAGCCCTTGGCATATCCACTGAAGTTT	1180	Best Local Similarity 81.8%; Pred. No. 1.1e-195;
Db	3844	AGGGCAAGTTTCCCTTAATCTCCACCTGAAGAGGCCCTCGGCTCGAGAGTGAAGTT	3903	Matches 783; Conservative 0; Mismatches 165; Indels
QY	1181	GCACAGCTCTGAGCTGCTCCGAGCCAGCTACTCTCTCC-----AGGAAGT	1231	9: Gaps 1
Db	3904	GCACGGCTCCAGCTACAGCCCATTTGAGGCAATACTGCTGCCAGTGGCAGCGGAACG	3963	
QY	1232	GGAGTGGGGTGAAGCTCCGGAGGAAGCTCTGGAGGAGTGGATTCTGTGCCGACAAAG	1291	
Db	3964	GGAGCGGAGTAGCAGCTCTGGAGGAGCTCTGGAGGAGTGGATTCTGTGCTGAGAG	4023	
QY	1292	CAGATGCCCTTACCCTGTGGCAGATGACAGAAATGCTTTTGGCAGTGCATCAATGAA	1351	
Db	4024	CCAAGGCCCTTACCCTGTGGCAGATGACAGAAATGCTTTTGGCAGTGCATCAATGAA	4083	
QY	1352	TCACATACAGCAGCATTTCAAGCAGGCTTGTGTTTGTATACAGCTGAATGCTGCA	1411	
Db	4084	TCACATACAGCAGCATTTCAAGCAGGCTTGTGTTTGTATACAGCTGAATGCTGCA	4143	
QY	1412	ACTGGCCATGAACCTAA	1428	
Db	4144	ACTGGCATAACTGA	4160	
RESULT 14				
AB025009		1188 bp mRNA linear	PRI 25-NOV-1999	
LOCUS		Homo sapiens TSAL902-S mRNA for novel member of chitinase family,		
DEFINITION		complete cds.		
ACCESSION		AB025009		
VERSION		AB025009.1 GI:6467178		
KEYWORDS		novel member of chitinase family; TSAL902-S.		
SOURCE		Homo sapiens cDNA to mRNA.		
ORGANISM		Homo sapiens		
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS		Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE		Saito,A., Ozaki,K., Fujiwara,T., Takahashi,E. and Tanigami,A.		
JOURNAL		Isolation and mapping of a human lung-specific gene, TSAL902,		
MEDLINE		encoding a novel chitinase family member		
REFERENCE		Gene 239 (2), 325-331 (1999)		
AUTHORS		20018184		
TITLE		2 (bases 1 to 1188)		
JOURNAL		Saito,A., Ozaki,K., Fujiwara,T., Takahashi,E. and Tanigami,A.		
MEDLINE		Submitted (16-MAR-1999) Akihiko Saito, Otsuka Pharmaceutical Co.,		
REFERENCE		Ltd., Otsuka GEN Research Institute; 463-10 Kagasuno Kawauchi-cho,		
AUTHORS		Tokushima, Tokushima 771-0192, Japan [E-mail:saito@otsuka.gr.jp,		
TITLE		Tel:81-88-665-2888, Fax:81-88-637-1035)		
JOURNAL		Location/Qualifiers		
MEDLINE		1. .1188		
REFERENCE		/organism="Homo sapiens"		
AUTHORS		/db_xref="taxon:9606"		
TITLE		1. .1188		
JOURNAL		/gene="TSAL902-S"		
MEDLINE		150. .1097		
REFERENCE		/gene="TSAL902-S"		
AUTHORS		/codon_start=1		
TITLE		/product="novel member of chitinase family"		
JOURNAL		/protein_id="BAAB86981.1"		
MEDLINE		/db_xref="GI:6467179"		
REFERENCE		/translation="MREAFQAEAKQINKPRLMVTAAVAGISNIQSYETPOLSYOLD		
AUTHORS		YIHVMYDLHGSWEVGTGENSLYKYPDTGSNAYLNDVYMYKDNKGAPAEKLIV		
TITLE		FPYGHNFILSNFTGIGTSGAGPAGPYAESGIWAYEYIEICTLKNAGTGWADP		
JOURNAL		QEVPIAYQGNVWYGVNDSKFDIAKQWLKHNFEGGAMVAIDLDITFTGFCNQGRPL		
MEDLINE		ISTLKKALQSACAPQPIEPIITAAPSGSGSGSGSGSGSGSGSGSGSGSGSGSGSG		
REFERENCE		YPVANRNNAFWHCVNGVTVQNCQAGLVFDTSCDCNWA"		
AUTHORS		280 a 337 c 309 g 262 t		
TITLE				
JOURNAL				
MEDLINE				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				

VERSION					
KEYWORDS	U56900.1	GI:1336165			
SOURCE	house mouse strain=CBA.				
ORGANISM	Mus musculus				
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.				
REFERENCE	1 (bases 1 to 994)				
AUTHORS	Shmelkov,S.V., Zinovjeva,M.V. and Belyavsky,A.V.				
TITLE	Mouse chitinase-related protein mRNA (MCRP), partial cds				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 994)				
AUTHORS	Shmelkov,S.V.				
TITLE	Direct Submission				
JOURNAL	Submitted (27-APR-1996) Sergey V. Shmelkov, Molecular Basis of Differentiation and Development, Engelhardt Institute of Molecular Biology, Vavilov 32, Moscow 117984, Russia				
FEATURES	Location/Qualifiers				

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 2, 2003, 22:31:11 ; Search time 349.54 Seconds
(without alignments)
9825.200 Million cell updates/sec

Title: US-10-004-219B-3

Perfect score: 1525

Sequence: 1 atggcagactactctctgt.....taaaattgtagccaaaca 1525

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_101002:*

- 1: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
- 2: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
- 3: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
- 4: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
- 5: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
- 6: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
- 7: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
- 8: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
- 9: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
- 10: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
- 11: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
- 12: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
- 13: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
- 14: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
- 15: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
- 16: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
- 17: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
- 18: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
- 19: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
- 20: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
- 21: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
- 22: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
- 23: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
- 24: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1042.6	68.4	1678	22 AAH42025	Disease treatment
2	999.2	65.5	1368	22 AAH42013	Disease treatment
3	854	56.0	1369	24 ABN59993	Novel human coding
4	714.6	46.9	1469	22 AAH42023	Disease treatment
5	426.8	28.0	1636	19 AAV10435	Human chitinase cl
6	426.8	28.0	1636	20 AAZ21847	MO-218 clone of hu
7	426.8	28.0	1636	22 AAD03759	Human chitinase cd
8	426.8	28.0	1636	24 ABL57380	Human chitinase cd
9	426.8	28.0	1643	18 AAT50833	Human chitinase cd

10	426.8	28.0	1713	18 AAT50834	Human chitinase cd
11	426.4	28.0	1768	18 AAT89181	Human chitotriosid
12	425.2	27.9	1656	19 AAV10436	Human chitinase cl
13	425.2	27.9	1856	20 AAZ21848	MO-13B clone of hu
14	425.2	27.9	1856	22 AAD03760	Human chitinase cd
15	425.2	27.9	1656	24 ABL57381	Human chitinase cd
16	406.4	26.6	1637	18 AAT89180	Human chitotriosid
17	362	23.7	2504	16 AAQ90443	Murine oviduct spe
18	350.8	23.0	2366	16 AAQ90444	Hamster oviduct sp
19	345.2	22.6	638	24 ABK11713	DNA encoding novel
20	344.8	22.6	746	22 AAK86635	Human digestive sy
21	343.2	22.5	1432	19 AAV21689	DNA encoding a hum
22	343.2	22.5	1433	19 AAV13925	Human cartilage gp
23	339.4	22.3	1994	16 AAQ90442	Bovine oviduct spe
24	338.8	22.2	1496	18 AAT97127	Human cartilage gl
25	335.6	22.0	1594	19 AAV21688	DNA encoding a hum
26	330.8	21.7	1526	19 AAV13926	Human cartilage gp
27	330.8	21.7	1594	19 AAV21687	DNA encoding a hum
28	327.6	21.5	1391	24 AAS62615	CDNA sequence #402
29	314.8	20.6	1681	16 AAQ85245	YKL-40 gene. Homo
30	314.8	20.6	1681	21 AA294901	Human cancer marke
31	307.4	20.2	1152	18 AAT99452	Bovine whey proteoi
32	296.6	19.4	2346	23 ABL03597	Drosophila melanog
33	264.4	17.3	17368	23 ABL03596	Drosophila melanog
34	259.8	17.0	1474	22 AAH23078	Osteoarthritis tis
35	210	13.8	966	9 AAN81756	Gene encoding poly
36	176.2	11.6	13497	23 ABL03557	Drosophila melanog
37	152.6	10.0	503	21 AAC00233	Human secreted pro
38	151.6	9.9	288	24 ABN76932	Human glycoprotein
39	146	9.6	2089	23 ABL29841	Drosophila melanog
40	142.2	9.3	3899	23 ABL19747	Drosophila melanog
41	140	9.2	2452	18 AAT62557	Manduca sexta larv
42	139.8	9.2	239	21 AAA42054	Human secreted exp
43	137.8	9.0	369	22 AAH42024	Disease treatment
44	136	8.9	1433	23 ABL15099	Drosophila melanog
45	132.6	8.7	1302	23 ABL19745	Drosophila melanog

ALIGNMENTS

RESULT 1
AAH42025
ID AAH42025 standard; DNA; 1678 BP.
AC AAH42025;
XX
XX
DT 24-AUG-2001 (first entry)
XX
DE Disease treatment related oligonucleotide SEQ ID NO: 16.
XX
XX Disease treatment; infection; chronic occlusive pulmonary disease;
KW bronchial asthma; ds.
XX
OS Homo sapiens.
XX
PN WO200136633-A1.
XX
PD 25-MAY-2001.
XX
PF 14-NOV-2000; 2000WO-JP08015.
XX
PR 15-NOV-1999; 99JP-0324467.
XX
PA (TAKE) TAKEDA CHEM IND LTD.
XX
PI Nakanishi A, Morita S;
XX
DR WPI; 2001-397791/42.
XX
PT New proteins, peptides and DNA for treatment of bronchial asthma,
XX chronic occlusive lung disease and infectious disease -
XX

Example 5; Page 107-108; 114pp; Japanese.

PS The present invention provides the sequence of a protein which can be
XX used in the treatment and prevention of infectious diseases. Inhibitors
CC of the protein can be used to treat bronchial asthma and chronic
CC occlusive pulmonary disease. The present sequence is an oligonucleotide
XX described in the exemplification of the invention.

SQ Sequence 1678 BP; 397 A; 497 C; 418 G; 366 T; 0 other;

Query Match 68.4%; Score 1042.6; DB 22; Length 1678;
Best Local Similarity 83.4%; Pred. No. 5.4e-307;
Matches 1199; Conservative 0; Mismatches 229; Indels 9; Gaps 1;
1 ATGCCAAGTACTTCTCGTCACAGGTCGGCTTCTGCTGAATGCTCAGCTGGGTCT 60
143 ATGACAAAGCTTATCTCTCAGAGTCTGCTTATATCTGATGCTGCGGTCT 202
61 GCCTACAACTGATATGCTATTTTACCAACTGGGCCAGATCGGCCAGCTCTGGGAGC 120
203 GCCTACAGCTGACATGCTACTTCCACCAACTGGGCCAGTACCGCCAGGCTGGGGCC 262
121 TTCAAGCTGATGACATTAACCCCTGCCTGTGTACTCACCTGATCTATGCTTGTGGG 180
263 TTCAAGCTGATGACATTAACCCCTGCCTGTGTACTCACCTGATCTATGCTTGTGGG 322
181 ATGCAGAACTAGATGATCACCACATAGAAATGAATGATCTTACTCTATAAGCTTTC 240
323 AGCAGAACACGAGATCACCACATAGAAATGAATGATCTTACTCTATAAGCTTTC 382
241 AATGACTTGAACAGGAGACAGCAAACTGAAACCCCTCTGCAATTTGGAGCTGGAAC 300
383 AATGGCTGAAATTAAGAACAGCCAGCTGAAACCTCTCTGCAATTTGGAGCTGGAAC 442
301 TTGGAGCTGCTCTTCTACTACATGCTTCTCAGTCTTCTCAGAACCCGACGCTTCATT 360
443 TTGGAGCTGCTCTTCTACTACATGCTTCTCAGTCTTCTCAGAACCCGACGCTTCATT 502
361 ACCTGAGTCAATAATTTCTGCTGATGATGGTGTGATGGTGGACCTGGAGTGGAA 420
503 ACCTGAGTCAATAATTTCTGCTGATGATGGTGTGATGGTGGACCTGGAGTGGAA 562
421 TACCCAGGCTCAGTGGGAGCCCTCTCAGGACAAAGATCTCTTACTGTCTCTGGTGAAG 480
563 TACCCAGGCTCAGTGGGAGCCCTCTCAGGACAAAGATCTCTTACTGTCTCTGGTGAAG 622
481 GAAATGCTGAAGCTTTGAGGAGGAGGCTATTGAGAGCAACAGGCTCAGACATGATGTT 540
623 GAAATGCTGAAGCTTTGAGGAGGAGGCTATTGAGAGCAACAGGCTCAGACATGATGTT 682
541 ACTGCTGCTGATGCTGGGAGTTCACCAATCCAGCTGGCTATGAGATCCCTGAACTT 600
683 ACTGCTGCTGATGCTGGGAGTTCACCAATCCAGCTGGCTATGAGATCCCTGAACTT 742
601 TCTAAGTACCTGATTTATCATCATGATGATGATGATGATGATGATGATGATGATGAT 660
743 TCAAGTACCTGATTTATCATCATGATGATGATGATGATGATGATGATGATGATGAT 802
661 TACACTGGGAGAAATAGTCTCTTTTACAAATCCCTTACTGAGACTGTAGCAATGCTTAC 720
803 TACACTGGGAGAAATAGTCTCTTTTACAAATCCCTTACTGAGACTGTAGCAATGCTTAC 862
721 CTCATGTGGATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
863 CTCATGTGGATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 922
781 ATTGTTGGATTTCCAGAGTATGACACACCTTCTATCTGAGAAACCCCTCTGATATGGA 840
923 ATGTTGGATTTCCAGAGTATGACACACCTTCTATCTGAGAAACCCCTCTGATATGGA 982
841 ATTGTTGGATTTCCAGAGTATGACACACCTTCTATCTGAGAAACCCCTCTGATATGGA 900
983 ATTGTTGGATTTCCAGAGTATGACACACCTTCTATCTGAGAAACCCCTCTGATATGGA 1042

QY 901 TGGSCCTACTATGAGATTTGCACCTTCTTGAGAAAGTGAGCCACTGAGGCTGGATGCC 960
1043 TGGSCCTACTACGAGATCTGTACCTTCTGAAATAGGACCACTCAGGATGGATGCC 1102
961 TCCCAAGAAGTGCCCTATGCTATAGGCCAACAGTGGCTTGGCTATGACATATCAAG 1020
1103 CCTCAGGAAGTGCCCTATGCTATCAGGCAATGTGTGGTGGCTATGACATCAAG 1162
1021 AGCTTCAGTGTAAAGGCTCAGTGGCTTAAGCAGAACAAATTTTGGAGTGCCATATCTGG 1080
1163 AGCTTCAGTGTAAAGGCTCAGTGGCTTAAGCAGAACAAATTTTGGAGTGCCATATCTGG 1222
1081 GCCATTGACCTTGAATGACTTCACTGGCTTCTTCTGATCAGGGAATAATTTCTCTGACT 1140
1223 GCCATTGATCTGGATGACTTCACTGGCACCTTCTGCAACAGGCAAGTTTCCCTTAATC 1282
1141 TCTACTTTGAACAAAGCCCTTGGCATATCCACTGAAGTGGCAGCTCTCTGAGCTGCT 1200
1283 TCCACCTGAAAGAGGCCCTCGGCTGCAGAGTCAAGTTGCAGGCTCCAGCTCAGGCC 1342
1201 TCCGAGCCAGTACTACTCTCC-----AGGAAGTGGAGTGGGTTGGAAGTCC 1251
1343 ATTGAGCAATTAAGTCTTCCAGTGGCAGGCGGAACGGGAGGAGTAGCAGCTCT 1402
1252 GGAGGAAGCTCTGGAGGAGTGGATTTCTGCCGACAAAGCAGATGGCCTCTACCCCTGTG 1311
1403 GGAGGAGCTCTGGAGGAGTGGATTTCTGCCGACAAAGCAGATGGCCTCTACCCCTGTG 1462
1312 GCAGATGACAAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1371
1463 GCAATTAACAAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1522
1372 CAAGCAGGCTGTGTTTGTATGACAGCTGTAATTTGTGCACTGGCCATGAACCTAA 1428
1523 CAGCGCGGCTGTGTTTGTGACACAGCTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1579

RESULT 2
AAH42013

ID AAH42013 standard; DNA; 1368 BP.

XX AAH42013;

XX 24-AUG-2001 (first entry)

DE Disease treatment related oligonucleotide SEQ ID NO: 3.

XX Disease treatment; infection; chronic occlusive pulmonary disease;
KW bronchial asthma; ds.

OS Homo sapiens.

PN WO200136633-A1.

XX 25-MAY-2001.

XX 14-NOV-2000; 2000WO-JP08015.

XX 15-NOV-1999; 99JP-0324467.

XX (TAKE) TAKEDA CHEM IND LTD.

XX Nakanishi A, Morita S;

DR WPI; 2001-397791/42.

XX New proteins, peptides and DNA for treatment of bronchial asthma,
PT chronic occlusive lung disease and infectious disease -

PS Claim 5; Page 100; 114pp; Japanese.

XX The present invention provides the sequence of a protein which can be

Db 9 ATGCCAGCTCATCTTCTGTACAGGCTGTGCAATCTCTCTGAAGCTACAGCTGGGATCT 68
QY 61 GCCTACATCTGATATGCTATTTTACCACACTGGGCCAGTAGTATCGCCAGGCTCTGGGAGC 120
Db 69 TCCCTACACAGCTGATGTGCTACTATACAGTGTGGCTAAGGACAGGCCAATAGAAGGAGT 128
QY 121 TTCAAGCCTGATGACATTAACCCCTGCTGTGTACTACCTGATGCTATATGCTCTTCTGCTGG 180
Db 129 TTCAACCTGCTATATGACCCCTGCTGTGTACTACCTGATGCTATATGCTCTTCTGCTGA 188
QY 181 ATGCAGAACATGAGATCACCAACATAGAATGGAATGATGTACTCTCTATAAAGCTTTTC 240
Db 189 ATGCAGATAATGATGATCATTACACATGAGCAAGACTTGCCTGACTATGAAGCAATG 248
QY 241 AATGACTTGAAGAACAGCAAGCAAACTGGAACCCCTCTGGCAATTTGGAGGCTGGAAAC 300
Db 249- ANTGGCTGAAGAACAGCAAGCAAGCTAGCTAAACTCTCTGGCCATTGGAGGATGGAAG 308
QY 301 TTGGAAGCTGCTCTTTTACCTACCTGCTTCCACTCTCAGAACCGCAGACCTTCATT 360
Db 309 TTGGAAGCTGCTCTTTTACCTGCTTCCACTCTCAGAACCGCAGACCTTCATT 368
QY 361 ACCTGAGTCAATATTTCTGGTCAGTATGGGTTTGTGAGTGGACTGGACCTGGAGTGGAA 420
Db 369 CAGTCAGTTATCAGATTCCTTCTGCTCAATATAACTTTGATGGCTCAACCTGGACTGGCAG 428
QY 421 TACCAGGCTCAGTGGAGGCTCTCTCAGCACAGCATCTCTCAGTCTGCTGGTGAAG 480
Db 429 TACCCTGGGCTCGAGGAAGCCCTCTAAGACAAACATCTCTCAGTCTGCTGGTGAAG 488
QY 481 GAAATGCTGAGCTTTTGGAGGAGGCTATTGAGAGCAACAGGCCCGAGACTGATGTT 540
Db 489 GAATGCTTAAGCTTTTGGAGGAGATCTGTGGAGAAAGACATTCGAAGGCTGCTACTC 548
QY 541 ACTGCTGCTGAGTGGGAGTTCCTTACAAATACCTCTGAGACTGGTAGCAATCCCTAC 720
Db 549 ACTTCCACAGGAGGAGGAATCATTGACGTAATCAAGTCTGGGTACAAGATCCCTGAAGTG 608
QY 601 TCTAGTACCTGGATTCATCCATGTCATGACATATGACCTCCATGGCTCTGGAGGAGGC 660
Db 609 TCTCAGTCTCTGACTATATTCAGTCAATGATATGATCTCCATGATCTCTTAAGGATGGC 668
QY 661 TACACTGGGAGATAGTCTCTTTACAAATACCTCTGAGACTGGTAGCAATCCCTAC 720
Db 669 TACACTGGAGAAATAGTCCCTCTATAAATCTCCATATGACATGGAAGAGTCTGAT 728
QY 721 CTCAATGGAATATGTCATGAATATTTGGAAGAACAAATGGAGGCCCGAGCTGAGAGCTC 780
Db 729 CTCAATGGAATCAATCATTTCTTACTGGAAGGACCATTGGAGCAGCTTCTGAGAAGCTC 788
QY 781 ATGTTGGATTCAGAGATGAGCACACACCTTCATCTCTGAGAAACCCCTCTGATATGGA 840
Db 789 ATTTGGGATTTCCAGCATATGGGCATACCTTTATCTCTGAGTGACCTTTCTAAGACTGA 848
QY 841 ATTTGGGCCCTTACCTCTGGGTGATGGCCCTGTCTGGCCCTATACACAGACAGGCTGGTTC 900
Db 849 ATTTGGGCCCTTACATTTAGTACTGGCCACCAGGAAGTACACAGATCAATCAGACTC 908
QY 901 TGGGCTTACTATGAGATTTGACCTTTCTGAGAAGTGGAGCCACTGAGGCTCTGGATGCC 960
Db 909 CTGGCTTACTATGAGTTTGTACATTTCTGAATGAAGGAGGACCCTGAGGCTCTGGATGCC 968
QY 961 TCCCAAGAGTCCCTATGCTTATAGCCACAGAGTGGCTTGGTATGACAAATATCAAG 1020
Db 969 CCCCAGGAATACCCCTATGCTATCAGGGTATGAGTGGGTTGGTTATGACAAATCTCAG 1028
QY 1021 AGCTTCACTGTTAAGGCTCAGTGGCTTAAGCAGAACAAATTTGGAGGTGCCATGATCTGG 1080
Db 1029 AGCTTCAAGTTGAAGGCTCAGTGGCTCAGGACACAAATTTAGGAGGTGCCGTGGTCTGG 1088
QY 1081 GCCATTGACCTTGAATGACTTCACTGGCTCTTCTGTGATCAGGGGAAATTTCTCTGACT 1140

Db 1089 CCCCTGGACATGATGACTTCAGTGGTCTTTCTGTCCACGAGACATTTCCCTCTGACA 1148
QY 1141 TCTACTTTGAACAAAGCCCTTGGCATATCCACTGAAGTGTGCACAGCTCCCTGA 1193
Db 1149 TCTACTTTAAAGGGAGATCTCAATATACAGTGCAGTGTGCAGGGCCCTTA 1201
RESULT 5
AAV10435
ID AAV10435 standard; cDNA; 1636 BP.
XX AAV10435;
AC AAV10435;
XX
DT 15-JUN-1998 (first entry)
XX Human chitinase clone MO-218 cDNA.
DE
XX Chitinase; human; fungal infection; immunogen; diagnosis; treatment;
KW Gaucher's disease; transgenic; detection; hybridisation; antifungal;
KW rheumatoid arthritis; overexpression; extracellular matrix; ss.
XX Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 2..1402
FT sig_peptide /tag= a
FT /tag= b
FT mat_peptide 65..1399
FT /tag= c
FT /product= chitinase
FT /note= "from clone MO-218"
XX
PN WO9747752-A1.
XX
PD 18-DEC-1997.
XX
PF 16-JUN-1997; 97WO-US10460.
XX
PR 14-JUN-1996; 96US-0663618.
XX
PA (ICOS-) ICOS CORP.
XX Gray PW;
PI
XX
XX
XX WPI: 1998-052316/05.
DR P-PSDB; AAW40259.
XX
PT Nucleic acids encoding human chitinase - useful as antifungal
agents, especially in combination with other antifungals
XX
PS Claim 3; Page 38-40; 63pp; English.
XX
CC This sequence encodes a novel human chitinase isolated from clone MO-218.
CC Chitinases are useful for treating or preventing fungal infection and
CC as immunogens for generating antibodies which are used to purify, detect
CC and quantify chitinases, e.g. for diagnosis of Gaucher's disease. The
CC nucleic acid sequence of the chitinase is also useful as a probe to
CC identify and isolate genomic DNA encoding chitinases or similar proteins,
CC or cells expressing them or to generate transgenic ('knockout') rodents.
CC It can also be used in hybridisation assays and to detect genetic
CC alterations in the chitinase gene related to disease. Agents that inhibit
CC this protein may be useful in treatment of Gaucher's disease and
CC rheumatoid arthritis, where overexpression of the protein can damage
CC the extracellular matrix. Chitinase also improves the activity of other
CC antifungal agents and may allow a reduction in the dose of such agents,
CC and thus of their side effects.
XX
SQ Sequence 1636 BP; 361 A; 491 C; 440 G; 344 T; 0 other;
Query Match 28.0%; Score 426.8; DB 19; Length 1636;
Best Local Similarity 62.1%; Pred. No. 3.4e-119;
Matches 693; Conservative 0; Mismatches 417; Indels 6; Gaps 1;

CC which are expected to be useful as products for treating fungal
 CC infections or for the development of such products. Human
 CC chitinase has a synergistic effect on the actions of other
 CC fungicides. It can be administered to improve the antifungal
 CC activity of a non-chitinase antifungal agent, especially
 CC amphotericin B or itraconazole, in the treatment of a fungal
 CC infection such as candidiasis, aspergillosis, coccidioidomycosis,
 CC blastomycosis, paracoccidioidomycosis, histoplasmosis,
 CC cryptococcosis, chromoblastomycosis, sporotrichosis, mucormycosis,
 CC dermatophytoses and Pneumocystis infections (all claimed). In
 CC particular, the fungal infection involves *Candida*, *Aspergillus*
 CC and/or *Cryptococcus* spp., whose growth is not effectively
 CC inhibited by contact with human chitinase alone.
 XX
 SQ Sequence 1636 BP; 361 A; 491 C; 440 G; 344 T; 0 other;

Query Match 28.0%; Score 426.8; DB 24; Length 1636;
 Best Local Similarity 62.1%; Pred. No. 3.4e-119;
 Matches 693; Conservative 0; Mismatches 417; Indels 6; Gaps 1;

QY 23 CAGGTCTGGCTCTTCTGCTGAATGCTCAGCTGGGTCTGCTTACAAATCTGATATCTATT 82
 DB 24 CAGGTCTGGCTCTTCTGCTGATGCTCCATGGGCTCTGCTGCAAACTGGTCTGCTACT 83

QY 83 TCACCAACTGGGCCAGTATCGGCAGCTCTGGGAGCTTCAAGCCTGATGACATTAACC 142
 DB 84 TCACCAACTGGGCCAGTATCGGCAGCTCTGGGAGCTTCAAGCCTGATGACATTAACC 143

QY 143 CCTGCTCTGCTACTACCTGATCTATGCTTGGTGGAGCTCAGAACATGATGATCACA 202
 DB 144 CAGGCTTTGGCACCACCTCATCTACGCTTTCGCTGGGATGATGACCAACCAGCTGGACA 203

QY 203 CCATAGAAATGGAATGATCTTACTCTATTAAGCTTTCAATGACCTTGAAGAACAGGACA 262
 DB 204 CCAGTGTGGAATGACGAGACTCTCTACGAGCTTCAATGGCTGGAAGATGATC 263

QY 263 GCAAACTGAAACCCCTCTCGCAATGAGGCTGGAATTTGGAATCTCTCTTTCACTA 322
 DB 264 CCAAGCTGAAGACCTGTGTAGCCATCGGAGGCTGGAATTTGGCACTCAGAAGTTCACAG 323

QY 323 CCATGGTTTCCACTTCTCAGAACCCGACGCTTCAATACCTCAGTCAATCAATTTTCG 382
 DB 324 ATATGGTAGCCGACGACCAACCTGCTGAGACTTTGTCACTCGGCCATCAGTTTCTGCG 383

QY 383 GTCACTATGGTGTGATGACTGACCTGGAGTACCCAGGCTCAGCTGGGAGCC 442
 DB 384 GCAATACAGCTTTGACGGCTTGACCTTGACTGGAGTACCCAGAGCAGGGAGCC 443

QY 443 CTCTCAGGACAAAGCATCTTTCACTGCTCTGGTGAAGAAATGCGTGAAGCTTTTGAGC 502
 DB 444 CTGCGGTAGAACAGGAGGCTTCAACACCTTGGTACAGGACTTGGCCAATGCTTCCAGC 503

QY 503 AGGAGGCTATTGAGAGCAACAGGCCAGCTGATGCTGTTACTGCTCTGCTAGCTGGTGA 562
 DB 504 AGGAAGCCGACAGCTCAGGAGGAGAACGCTTCTTCTGAGTGACGCGGTTCAGCTGGGC 563

QY 563 TTTCACACATCCAGCTGGCTATGAGATCCCTGTAAGTACCTGGAATTTTCATCC 622
 DB 564 AGACCTATGTGATGCTGATGACGAGTGGACAAATCGCCAGAACCTGGATTTGTCA 623

QY 623 ATGTCATGACATATGACCTCCATGGCTCTCTGGAGGGCTACACTGGGAGAAATAGTCCTC 682
 DB 624 ACCTATGGCTTACGACTTCCATGCTCTTGGGAGAGGTACGCGACATACAGCCGCC 683

QY 683 TTTCAAAATACCTTACTGAGACTGGTACGAATGCCCTACCTCAATGTGATATGTCATGA 742
 DB 684 TCTACAGAGGCAAGAGAGAGTGGTGACGAGCAGCCCTCAACGTGGATGCTGTGTCG 743

QY 743 ACTATTGGAAGAACACCCAGCTGAGAGCTCATTTGTTGATTTCCAGAGATG 802
 DB 744 AACAGTGGCTGCAGAGAGGGGCCCTCGCAGCAAGCTGATCTCTGGCATGCTTACTAGC 803

QY 803 GACACACCTTCACTCTGAGAAACCCCTCTGTATAATGGAATTTGGTGGCCCTACCTCTGGTG 862

DB 804 GACGTCCTTTCACACTGGCCCTCTCATCAGACACCAGAGTGGGGCCCCAGCCACAGGT 863
 QY 863 ATGGCCCTGCTGGGCCCTTATACGACAGGCTGGCTTCTGGCCCTACTATGAGATTTGCA 922
 DB 864 CTGGCATTCCAGGCCCTTTCACCAAGAGGAGGAGTCTGGCCCTACTATGAAGTCTGCT 923
 QY 923 CCTTTCTGAGAAGTGGAGCCACTGAGGTCTGGGATGCTCCCAAGAAGTGGCCCTATGCT 982
 DB 924 CCT-----GGAAGGGGGCCACCAACAGAGAAATCAGGATCAGAAGTGGCCCTACATCT 977
 QY 983 ATAAGGCCAACAGAGTGGCTTGGCTATGACATATCAAGAGCTTCAAGTGTTAAGGCTCAGT 1042
 DB 978 TCCGGGACACCACTGCTGGGCTTTGATGATGTGAGAGCTTCAAAAACCAAGGTCAGCT 1037
 QY 1043 GGCTTAAGCAGAACAAATTTTGGAGGTGCCATGATCTGGGCCATTGACCTTGATGACTTCA 1102
 DB 1038 ATCTGAACAGAAAGGAGTGGCGGGGCCATGCTGGGCACTGGCACTAGATGACTTTG 1097
 QY 1103 CTGGCTCTTCTGTGATCAGGGAATTTTCCTCTCA 1138
 DB 1098 CCGGCTTCTCTGTGCAACCAGGCGCGCATACCCCTCA 1133

RESULT 9
 AAT50833
 ID AAT50833 standard; cDNA; 1643 BP.
 XX AAT50833;
 AC AAT50833;
 XX
 DT 24-MAR-1997 (first entry)
 XX Human chitinase cDNA clone chi.50.
 DE
 XX
 KW Chitinase; chitotriosidase; chitin; infectious disease;
 KW gene therapy; vaccine; lysosomal lipodosis; Gaucher's disease;
 KW leishmaniasis; sarcoidosis; X-linked adrenoleukodystrophy;
 KW multiple sclerosis; drug delivery; cosmetics; food; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 CDS 13..1413
 FT /*tag= a
 FT sig_peptide 13..75
 FT /*tag= b
 FT mat_peptide 76..1410
 FT /*tag= c
 XX
 PN WO9640940-A2.
 XX
 PD 19-DEC-1996.
 XX
 PF 06-JUN-1996; 96WO-NL00225.
 XX
 PR 07-JUN-1995; 95US-0486839.
 XX
 XX (UNAM) UNIV AMSTERDAM.
 XX
 PI Aerts JMWG;
 XX
 DR WPI; 1997-118698/11.
 DR P-PSDB; AAW08584.
 XX
 PT New human chitinase - used to treat or prevent infection by
 PT chitin-contg. pathogens, in diagnosis and as additives to cosmetics,
 PT foods, implants etc.
 XX
 PS Claim 2; Page 39-40; 58pp; English.
 XX
 CC A cDNA clone (AAT50833), designated chi.50, codes for a 50 kDa human
 CC chitinase (AAW08584) that is stable to many proteases, active at pH 3-
 CC 8 and up to 50 deg, and stable in the circulation. Clones chi.50

Db	875	CTGGCACTCAGAGCCCTTCACCAAGGAAGAGGATGCTGGCCCTACTATGAAGTCTGCT	934
Qy	923	CCTTTCTCAGAAGTGGAGCCCACTGAGTCTCTGGATGCTCCCAAGAAAGTGCCCTATGCCT	982
Db	935	CCT-----GGAAGGGGCCACCAACAGAGAATCCAGGATCAGAAGGTGCCCTACATCT	988
Qy	983	ATAAGGCCAAGAGTGGCTTGGCTATGACAATATCAAGAGCTTCAGTGTTAAGGCTCAGT	1042
Db	989	TCCGGGACAACCACTGGGTGGCTTTTGATGATGTGGAGAGCTTCAAAACCAAGGTCAGCT	1048
Qy	1043	GGCTTAAGCAGAACAAATTTTGGAGGTGCCATGATCTGGGCCATTGACCTTGATGACTTCA	1102
Db	1049	ATCTGAAGCAGAAGGACTGGCGGGGCCATGGTCTGGGCACCTGGACTTAGATGACTTTG	1108
Qy	1103	CTGCTCTTCTCTGTGATCAGGGAAAATTTCTCTGA	1138
Db	1109	CCGGCTTCTCTGTCAACACGAGGGCGGATACCCCTCA	1144
RESULT 10			
ID	AAT50834 standard; cDNA; 1713 BP.		
XX	AAT50834;		
AC	24-MAR-1997 (first entry)		
DT	Human chitinase cDNA clone chi.39.		
XX	Chitinase; chitotriosidase; chitin; infectious disease;		
KW	gene therapy; vaccine; lysosomal lipidosis; Gaucher's disease;		
KW	leishmaniasis; sarcoidosis; X-linked adrenoleukodystrophy;		
KW	multiple sclerosis; drug delivery; cosmetics; food; ss.		
OS	Homo sapiens.		
XX	Key		
FH	Location/Qualifiers		
FT	13..1176		
FT	/*tag= a		
FT	13..75		
FT	/*tag= b		
FT	76..1173		
FT	/*tag= c		
XX	WO9640940-A2.		
PN	19-DEC-1996.		
XX	PD		
XX	06-JUN-1996; 96WO-NL00225.		
PF	07-JUN-1995; 95US-0486839.		
XX	(UNAM) UNIV AMSTERDAM.		
PA	Aerts JMFG;		
XX	WPI; 1997-118698/11.		
DR	P-PSDB; AAW08585.		
XX	New human chitinase - used to treat or prevent infection by		
PT	chitin-contg. pathogens, in diagnosis and as additives to cosmetics,		
FT	foods, implants etc.		
XX	Claim 2; Page 42-43; 58pp; English.		
XX	A cDNA clone (AAT50834), designated chi.39, codes for a 39 kDa human		
CC	chitinase (AAW08584) that is stable to many proteases, active at pH 3-		
CC	8 and up to 50 deg, and stable in the circulation. Clones chi.39		
CC	and chi.50 (see also AAT50833) were isolated from a human macrophage		
CC	cDNA library using as probe a partial clone obtd. using primers		
CC	(see also AAT50835-36) based on a chitotriosidase purified from a		
CC	type 1 Gaucher disease patient. The 2 cDNA clones are the result		
CC	of alternative splicing of RNA. Chitinase nucleic acid can be used		

CC for large-scale prodn. of recombinant human chitinases, or can be
CC incorporated into a gene therapy vector to treat or prevent
CC infection by chitin-contg. pathogens.

XX Sequence 1713 BP; 380 A; 503 C; 465 G; 365 T; 0 other;

Query Match 28.0%; Score 426.8; DB 18; Length 1713;
Best Local Similarity 62.1%; Pred. No. 3.5e-119;
Matches 693; Conservative 0; Mismatches 417; Indels 6; Gaps 1;

```
QY 23 CAGGTCTGGCTCTCTGCTGAATGCTCAGCTGGGGTCTGCGCTACAACTGATATGCTATT 82
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 35 CAGGTCTGCTGCTCTGCTGATGCTCAGCTGGGGTCTGCTCAAACTGCTGCTACT 94
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 83 TCACCAACTGGGCCAGTATCGGCCAGCTTGGGAGCTTCAAGCTGATGACATTAACC 142
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 95 TCACCAACTGGGCCAGTATCGGCCAGCTTGGGAGCTTGGCTTCTGCCAAGACTTGGACC 154
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 143 CCGTCTGCTGCTACTGCTGCTATGCTATGCTTGGTGGATGCGAGCAAACTGAGATCACCA 202
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 155 CCGCTTTGACCCACCTATCTAGCTTGGCTTGGCTGCGATGACCAACACCCAGCTGAGCA 214
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 203 CCATAGATGAATGATGTTACTCTCTATTAAGCTTTCATGACTTGAATAACAGAGACA 262
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 215 CCAGTGTGAATGACGAGACTCTCTACAGGAGTTCATGGCTGAAGAGATGAATC 274
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 263 GCAAACTGAAACCTCTCTGCAATTTGGAGCTGGAATTTGGAACCTCTCTTCACTA 322
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 275 CCAAGCTGAGACCTGTTAGCCATCGGAGGCTGGAATTTGGCACTCAGAACTTCCAG 334
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 323 CCATGTTTCCACTTCTCAGAACCCAGACCTTCATTACCTCAGTCACTAAATTTCTGC 382
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 335 ATATGTTAGCCAGGCCAACCAACCTGTCAGACTTTGTCACTCGGCCATCAGTTTCTGC 394
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 393 GTCAGTATGGTTGATGAGCTGACCTGGAGCTGGGAATACCCAGGCTCAGTGGGAGCC 442
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 395 GCAATAGAGCTTTTGACGGCTTGACCTTGGAGTACCCAGGAAGCCAGGAGGCC 454
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 443 CTCCTCAGCAAGCATCTTCTACTGCTCTGCTGAGGAATGCTGAAGCTTTTGAGC 502
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 455 CTGCGGTAGACAGGAGGCTTCAACACCTGTTACAGACTTGGCCAACTGCTTCCAGC 514
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 503 AGGAGCTATTGAGAGCAACAGCCAGCTGATGTTACTGCTGCTAGTGTGGGA 562
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 515 AGGAAGCCAGACCTCAGGAAGAACGCTTCTCTGAGTGCACGCTTCCAGCTGGC 574
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 563 TTTCCAACTCCAGCTGGCTATGAGATCCCTGAACTTTCTAAGTACCTGGAATTCATCC 622
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 575 AGACCTATGCTGATGCTGAGTGGAGTGGACAAATCGCCAGAACTGGATTTGTCA 634
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 623 ATGTCATGACATATGACCTCGATGCTCTGAGGAGGCTACACTGGGAGATACCTCCTC 682
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 635 ACCTTATGGCTACGCTTTCATGCTTGGGAGAGGTACGGGAGATACAGCCCCC 694
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 693 TTTCAAAATACCTACTGAGACTGCTAGCTGCTCAATGTTGATGTCATGA 742
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 695 TCTACAGAGCAAGAGAGAGTGTGTCAGAGCCAGCTCAACGTGATGCTGTGTC 754
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 743 ACTATTGAGAGCAATGAGCCCGAGCTGAGAGCTCATTTGTTGGATTCCAGATATG 802
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 755 AACAGCTGTGAGAGGGGAGCCCTGTCAGCAAGCTGATCTTGGCATGCTTACCTACG 814
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 803 GACACACTTCACTCTGAGAACCCCTCTGATAATGGAATGTTGCTCCCTACCTCTGGTG 862
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 815 GAGGCTCTTCACTGCGCTCTCATCAGACACAGAGTGGGGGCCCCAGCAGAGGT 874
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 863 ATGGCTGCTGCGCCCTATACACAGAGGTGCTGCTGCGCTTACTATGAGATTGCA 922
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 875 CTGGCACTCAGGCCCTTCCACAGAGGAGGAGTCTGCTGCTTACTATGAGTCTGCT 934
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 923 CCTTCTCAGAGTGGAGCCACTAGTGTCTGGGATGCTCCCAAGAGTGCCTATGCT 982
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 935 CCT-----GGAAGGGGGCCCAACACAGAGAAATCCAGGATCAGAAAGTGCCTACATCT 988
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
QY 983 ATAAGGCCAACGAGTGGCTTGGCTATGACAATATCAAGAGCTTCAAGTGTAAAGCTCAGT 1042
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 989 TCCGGGACACAGTGGTGGCTTGGCTTGGATGATGCTGGAGCTTCAAAACCAAGTCAAGT 1048
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1043 GGCCTTAAGCAGAACAAATTTTGGAGTGGCTGATGATGCGGCCATGACCTTGTATGACTTCA 1102
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1049 ATCTGAACAGCAAGGAGTGGCGGGGCGCATGCTGCGCACTGAGACTTAGATGACTTTG 1108
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1103 CTGGCTCTTCTGCTGATCAGGGAATTTTCTCTCA 1138
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1109 CCGGCTTCTGCAACCAAGGCGCATACCCCTCA 1144
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 11
AAT89181
ID AAT89181 standard; cDNA; 1768 BP.
XX AAT89181;
AC AC
DT 27-APR-1998 (first entry)
XX Human chitotriosidase variant cDNA.
DE Human chitotriosidase variant cDNA.
XX
XX Chitotriosidase; tissue remodelling disorder; diagnosis; therapy;
KW rheumatoid arthritis; atherosclerosis; human; ss.
XX Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 124..1524
FT tag= a
XX
XX WO9736917-A1.
XX
XX 09-OCT-1997.
XX
XX 21-MAR-1997; 97WO-US05072.
XX
XX 29-MAR-1996; 96US-0014295.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX (SMIK ) SMITHKLINE BEECHAM CORP.
XX
XX Chopra A, Gentz R, Kirkpatrick RB, Ni J, Thotakura NR;
XX
XX WPI; 1997-503041/46.
XX P-PSDB; AAW31498.
XX
XX New isolated human chitotriosidase gene - used to develop products
XX for the diagnosis and treatment of tissue remodeling disorders, e.g.
XX rheumatoid arthritis
XX
XX Claim 1; Page 21-22; 34pp; English.
XX
XX This human cDNA sequence encodes chitotriosidase (see AAW31498). It
XX is a preferred variant of the chitotriosidase cDNA sequence given
XX in AAT89180. Also claimed are: (1) a method of diagnosing a tissue
XX remodelling disorder related to expression of a mutated
XX chitotriosidase protein in a host comprising carrying out nucleic
XX acid amplification; and (2) a method of detecting altered
XX expression of a chitotriosidase protein in a host comprising
XX contacting a bodily sample with an antibody. Chitotriosidase can
XX degrade extracellular matrix substrates with a similar carbohydrate
XX structure to chitin. It can be used to develop products which can
XX be used in the diagnosis and treatment of tissue remodelling
XX disorders such as rheumatoid arthritis or atherosclerosis.
XX Chitotriosidase nucleic acids can be used in the recombinant
XX production of the enzyme, and as probes or primers in diagnostic
XX assays.
XX
XX Sequence 1768 BP; 401 A; 529 C; 474 G; 363 T; 1 other;
```

Query Match 28.0%; Score 426.4; DB 18; Length 1769;
Best Local Similarity 62.0%; Pred. No. 4.7e-119;
Matches 692; Conservative 1; Mismatches 417; Indels 6; Gaps 1;

QY 23 CAGGCTGGCTCTCTGCTGAATGCTCAGCTGGGGTCTGCTCAATCTGATATGATAT 82
DB 146 CAGGTTTCATGCTGCTGATGATCCATGGGGCTCTGCTGCAAACTGGTCTGCTACT 205
QY 83 TCACCAACTGGGCCAGGATTCGGCCAGGCTTCAGAGCTTCATGATGATTAAC 142
DB 206 TCACCAACTGGGCCAGGATTCAGACAGGGAGGCTGCTTCCTGCCAAGGACTTGGACC 265
QY 143 CTTGCTGTGCTACTACCTGATCTATGCTTCTGCTGGATGTCAGAACATGATGATCACA 202
DB 266 CCAGCTTTCACCCACCTCATCTAGCCTTCGCTGCTGATGACCAACACCCAGCTGAGCA 325
QY 203 CCATAGATGGAATGATGTTACTCTCTATTAAGCTTTCATGATGATTAACAGCA 262
DB 326 CCAGCTGAGTGAATGAGAGACTCTCTACGAGGATTAATGCGCTGGAAGATGAATC 385
QY 263 GCAAACTGAAACCCCTCTGGAATGAGGCTGGAATTTGGAATGCTCTCTTCACTA 322
DB 386 CCAAGCTGAAGACCTTTAGCCATCGAGGCTGGAATTCRCACCTCAGAAGTTCACAG 445
QY 323 CCATGTTTCCACTTCTCAGAACCCGACACCTTCATTAACCTCAGTCATCAATTTCTGC 382
DB 446 ATATGTTAGCCAGCGGCAACACCGCTCAGACCTTTCTCAACTCGGCCATCAGGTTCTGC 505
QY 383 GTCAGTATGGTTTGTATGACTGGACCTGGACTGGGAATACCCAGGCTCAGCTGGAGCC 442
DB 506 GCAATACAGCTTTGACGGCTTGACCTTGACTGGGATACCCAGGAGCAAGCCAGGGAGCC 565
QY 443 CTCTCAGGACAGCACTCTTTCACCTGCTGCTGGAAGAAATGCTGAAGCTTTTGAGC 502
DB 566 CTGCGCTAGCAGAGGAGCTTTCACAACTGTCAGGACTTGGCCATGCTTCCAGC 625
QY 503 AGAGGCTATTGAGACAACAGGCCAGACTGATGTTACTGCTGTGCTGAGCTGGTGGGA 562
DB 626 AGGAGCCCAAGACCTCAGGGAAGAGAGCTTCTGAGTGCAGCGGTTCCAGCTGGGC 685
QY 563 TTTCCAACTCAGGCTGCTGATGATCCCTGAACCTTCTAGTACCTGGATTTTCATCC 622
DB 686 AGACCTATGATGCTGGATGAGGAGTGAGACAAATCGCCCAAGACCTGGATTTGTCA 745
QY 623 ATGTCATGATATGATACCTCCCTGCTGCTGGAGGCTACACTGGGAGAAATAGTCCTC 682
DB 746 ACCTTATGCTACGACTTCCATGCTCTTGGGAGAGGTCACGGGACATACAGCCCCC 805
QY 683 TTTACAAATACCTTACTGAGCTGTGATGCAATGCTTACCTCAATGTGGATTAATGATGA 742
DB 806 TCTACAGAGGCAAGAGAGAGTGTGTCAGCAGCCCAAGCTCAACGTGGATGCTGTGTC 865
QY 743 ACTATTGGAAGAACATGAGCCCAAGCTGAGAGCTCATTTGTTGATTCACAGAGTATG 802
DB 866 ACAGTGGCTGAGAGGGGACCCCTGCGAGAGGCTGATCTTGGCATGCTTACCTACG 925
QY 803 GACACACCTTCTCTGAGAAACCCCTCTGATAATGGAATGCTGCTGCTTCTGCTG 862
DB 926 GAGCTCTTCACTGGCTCTCTCATCAGACACAGAGTGGGGGCCACGACAGGT 985
QY 863 ATGGCCCTCTGGGCTTATACAGACAGCTGGGTTGCTGGGCTTACTATGATGATTTGA 922
DB 986 CTGGCACTCCAGGCCCTTTCACCAAGAGAGGAGTGTGCTTACTATGAATGCTGCT 1045
QY 923 CTTTCTGAGAGTGGAGCCACTGAGGCTGGATGCTCCCAAGAGTGGCCCTATGCT 982
DB 1046 CCT-----GGAAGGGGGCCCAACACAGAGATCCAGGATCAGAGGTGCTCATCT 1099
QY 983 ATAAGGCCACAGGTGCTTGGCTATGACAATTAAGAGCTTCAGTTAAGGCTCAGT 1042
DB 1100 TCCGGGACACACAGTGGGCTTTGATGATGTGGAGAGCTTCAAAACCAAGGTGAGT 1159
QY 1043 GGCCTTAAGCAGAACAAATTTTGGAGGTGCTGATCTGGGCCATGACCTTGTGATGATCA 1102

DB 1160 ATCTAAGCAGAGGAGGACTGGGGGCGCATGCTGGGCACTGGACTTAGATGACTTTG 1219
QY 1103 CTGGCTCTTCTGCTGATCAGGGAATTTCTCTGA 1138
DB 1220 CCGGCTTCTCTGCAACACGAGGCGCATACCCCTCA 1255

RESULT 12
AAV10436
ID AAV10436 standard; cDNA; 1656 BP.
XX AAV10436;
XX AC
XX AC
XX 15-JUN-1998 (first entry)
XX Human chitinase clone MO-13B cDNA.
DE Chitinase; human; fungal infection; immunogen; diagnosis; treatment;
KW Gaucher's disease; transgenic; detection; hybridisation; antifungal;
KW rheumatoid arthritis; overexpression; extracellular matrix; ss.
XX Homo sapiens.
XX Key Location/Qualifiers
FH CDS 27..1427
FT /*tag= a
FT sig_peptide 27..89
FT /*tag= b
FT mat_peptide 90..1424
FT /*tag= c
FT /product= chitinase
FT /note= *from clone MO-13B*
XX WO9747752-A1.
XX 18-DEC-1997.
XX 16-JUN-1997; 97WO-US10460.
XX 14-JUN-1996; 96US-0663618.
XX (ICOS-) ICOS CORP.
XX Gray PW;
XX PI
XX WPI: 1998-052316/05.
XX P-PSDB; AAW40260.
XX Nucleic acids encoding human chitinase - useful as antifungal
PT agents, especially in combination with other antifungals
XX Claim 9; Page 42-44; 63pp; English.
XX This sequence encodes a novel human chitinase isolated from clone MO-13B.
CC Chitinases are useful for treating or preventing fungal infection and
CC as immunogens for generating antibodies which are used to purify, detect
CC and quantify chitinases, e.g. for diagnosis of Gaucher's disease. The
CC nucleic acid sequence of the chitinase is also useful as a probe to
CC identify and isolate genomic DNA encoding chitinases or similar proteins,
CC or cells expressing them or to generate transgenic ('knockout') rodents.
CC It can also be used in hybridisation assays and to detect genetic
CC alterations in the chitinase gene related to disease. Agents that inhibit
CC this protein may be useful in treatment of Gaucher's disease and
CC rheumatoid arthritis, where overexpression of the protein can damage
CC the extracellular matrix. Chitinase also improves the activity of other
CC antifungal agents and may allow a reduction in the dose of such agents,
CC and thus of their side effects.
XX Sequence 1656 BP; 365 A; 497 C; 447 G; 347 T; 0 other;
XX Query Match 27.9%; Score 425.2; DB 19; Length 1656;
XX Best Local Similarity 62.0%; Pred. No. 1.1e-118;

Matches	692;	Conservative	0;	Mismatches	418;	Indels	6;	Gaps	1;
QY	23	CAGGTCTGGCTCTTCTGCTGAATGCTCTCAGCTGGGCTGTGCTACAAATCTGATATSCATT	82						
DB	49	CAGGTTTCACTGGCTGCTGATGCCATCCATGGGCTCTGCTGCAAACTGCTGCTGCTACT	108						
QY	83	TCACCAACTGGGCGCCAGTATCGGCAGGCTCTGGGAGGCTTCAAGCCCTGATGACATTAACC	142						
DB	109	TCACCAACTGGGCGCCAGTACAGACAGGGGAGGCTGCTTCTGCCCAAGGACTTGGACC	168						
QY	143	CTGCTCTGTACTCACTCATCTATGCCCTTTCTGGGATGCGAACAATGAGATCACCA	202						
DB	169	CGAGCTTTCACCCACCTCATCTACGCCCTTCGCTGGCATGACCAACCACAGCTGAGCA	228						
QY	203	CCATAGAAATGGAATGATTTACTCTCTATAAGCTTTTCAATGACTTTGAAAACACAGAACCA	262						
DB	229	CCACTGAGTGAATGACGAGACTCTCTACCGAGGATTCATATGGCTGAAGAAGATGAATC	288						
QY	263	GCAAACTGAAAACCTCTCTGGCAATTTGGAGGCTTGGAACTTTTGGAACTTGCCTTTTCACTA	322						
DB	289	CCAAGCTGAAGACCTGTTTACCCATCGAGGCTTGGAAATTTTACGACACTCAGAAAGTTACAG	348						
QY	323	CCATGTTTTCCACTTCTCAGAACGCCAGACCTTTCATTAACCTCAGTCATCAAAATTTCTGC	382						
DB	349	ATATGTTAGCAGCGCCCAACCGTCAGACCTTTGTCAACTCTGGCCACTCAGGTTTCTGC	408						
QY	383	GTCAAGTATGGTTTTGATGGACTGGACCTGGACTGGGAATACCCAGGCTCAGCTGGGAGCC	442						
DB	409	GCAATACACTTTGACGGCTTGACCTTTCCTGCTGGAGTACCCAGGAAGCCAGGGGAGCC	468						
QY	443	CTCCTCAGGACACAGCATCTCTTCACTGCTGTGGTGAAGAAATCGTGAAGCTTTTGGAGC	502						
DB	469	CTGCGGTAGACAAGGAGCGCTTTCACACCCCTGGTACAGGACTTGGCCCAATGCCTTCCAGC	528						
QY	503	AGGAGGCTATTGAGAGCAACAGGCCACAGACTGATGTTACTGCTGCTGTAGCTGGTGGGA	562						
DB	529	AGGAAGCCACAGACCTTCAGGGAAGGAACGCCCTTCTTCTGAGTGCAGCGGTTCCAGCTGGC	588						
QY	563	TTTTCAACATCCAGGCTGGCTATGAGATCCCTGAACCTTCTTAAGTACTCGATTTTCAATCC	622						
DB	589	AGACCTATGTGGATGCTGGATACGAGGTGACAAATCGCCAGAACCTGGATTTTGTCA	648						
QY	623	ATGTCATGACATATGACCTCCATGGCTCTCTGGGAGGCTACACTGGGGAGAATAGTCCCT	682						
DB	649	ACCTTATGGCTACGACTTCCATGGCTCTTGGGAGAGGTCACGGGACATAACAGCCCCC	708						
QY	683	TTTTACAATACCCCTACTGACACTGGTAGCAATGCCCTACCTCAATGTGGATTATGTCAATGA	742						
DB	709	TCTACAGAGGCAAGAGAGAGTGGTCAGACGCCAGCCTCAAGCTGGATGCTGCTGTGC	768						
QY	743	ACTATTGGAAGACAATGGAGCCCACTGAGAGGCTCATTTGTTGGATTCCACAGATATG	802						
DB	769	AACAGTGGCTGCAGAGGGGAGCCCTGCCAGCAAGCTGATTCCTTGGCATGCTTACCTACG	828						
QY	803	GACACACCTTTCATCCTTGAGAAACCCCTCTGATTAATGGAATTTGGTGCCTTACCTTGGTG	862						
DB	829	GACGCTCTTTCACACTGGCTCTCTCATCAGACACAGAGTGGGGGCCCCCAGCCACAGGT	888						
QY	863	ATGGCCCTGTGGGCCCTATACCNAGACAGGCTGGGTTCTGGGCCCTACTATGAGATTTGCA	922						
DB	889	CTGGCACTCCAGGCCCTTCCACCAAGGAAGGAGGATGCTGGCCCTACTATGAAGTGTGCT	948						
QY	923	CTTTTCTGGAAGTGGAGCCACTGAGTCTTGGGATGCCCTCCCAAGAAGTGCCTATGCCT	982						
DB	949	CCCT-----GGAAGGGGGCCCAACAGAGATCCAGAGTTCAGAGGTGGCCCTACATCT	1002						
QY	983	ATAAGGCCAACGATGGCTTTGGCTATGACAATATCAAGAGCTTCAGTGTGTTAAGGCTCAGT	1042						
DB	1003	TCCGGGACACCACTGGGTGGCTTTGATGATGTGGAGAGCTTCAAAACCAAGGTCACT	1062						
QY	1043	GGCTTAAGCAGACAAATTTTGGAGGTGCCATGATCTGGGCCATTTGACCTTTGATGACTTCA	1102						
DB	1063	ATCTGAAGCAGAGGAGGACTGGGGGGGCCATGGTCTGGGCACCTGCACTTTAGATCACTTTG	1122						

Qy	1103	CTGCTCTTTCTGTGATCAGGAAATTTCTCTGA	1138
Db	1123	CGGCTTCTCTGTCAACCGGCGGATACCCCTCA	1158
RESULT 13			
AAZ21848	AAZ21848 standard; DNA; 1656 BP.		
XX	AC	AAZ21848;	
XX	DT	10-DEC-1999 (first entry)	
XX	XX	MO-13B clone of human Chitinase, with noncoding 5'/3' regions.	
DE	XX	chitin; fungal infection; immunocompromised; AIDS; chemotherapy;	
KW	KW	organ transplant; parasite; chitin-binding; allele; vector;	
KW	KW	truncated protein; ds.	
XX	XX		
OS	XX	Homo sapiens.	
XX	XX		
FH	FT	Key	Location/Qualifiers
FT	FT	CDS	27..1427
FT	FT		/*tag= a
FT	FT		/product= Human_Chitinase
FT	FT	sig_peptide	27..89
FT	FT		/*tag= b
FT	FT		/note= "Signal peptide"
FT	FT	mat_peptide	90..1427
FT	FT		/*tag= C
FT	FT		/note= "Mature peptide"
XX	XX		
PN	XX	W09946390-A1.	
XX	XX		
PD	XX	16-SEP-1999.	
XX	XX		
PF	XX	12-MAR-1999; 99WO-US05343.	
XX	XX		
PR	XX	12-MAR-1998; 98US-0039198.	
XX	XX		
PA	XX	(ICOS-) ICOS CORP.	
XX	XX		
PI	XX	Gray PW, Tjoelker LW;	
XX	XX		
DR	XX	WPI; 1999-551417/46.	
DR	XX	P-PSDB; AAY42426.	
XX	XX		
PT	PT	Novel chitin-binding fragments of human chitinase used to treat fungal infections in animals	
XX	XX		
PS	XX	Example 1; Page 59-62; 83pp; English.	
XX	XX		
CC	CC	This is the nucleotide sequence of an allelic form of the human chitinase enzyme, which is capable of degrading Chitin (a linear homo polymer of beta-1,4-linked N-acetylglucosamine residues).	
CC	CC	Chitinase fragments can be used to screen for proteins or other molecules that specifically bind to the chitin-binding domain of human chitinase or that modulate its activity. These compounds are useful for immunization, as well as for purifying chitinase, as well as for detection and quantification of chitinase. Polynucleotide fragments of the invention are useful as a source of probes and primers, and to express the proteins recombinantly. The chitinase fragments, when conjugated to antifungal compounds, are used to treat animals, especially humans, infected with chitin-containing parasites such as fungi. Fungal infection treated include candidiasis, aspergillosis, coccidioidomycosis, blastomycosis, paracoccidioidomycosis, mucormycosis, histoplasmosis, cryptococcosis, chromoblastomycosis, sporotrichosis, and dermatophytoses.	
CC	CC	Chitin can be degraded by the enzyme chitinase. Use of whole chitinase protein for treating infections, especially fungal infections, is problematic. In view of the increasing incidents of life-threatening fungal infection in e.g. immunocompromised individuals, there exists a	

methods for the recombinant production of human chitinase products, which are expected to be useful as products for treating fungal infections or for the development of such products. Human chitinase has a synergistic effect on the actions of other fungicides. It can be administered to improve the antifungal activity of a non-chitinase antifungal agent, especially amphotericin B or itraconazole, in the treatment of a fungal infection such as candidiasis, aspergillosis, coccidioidomycosis, blastomycosis, paracoccidioidomycosis, histoplasmosis, cryptococcosis, chromoblastomycosis, sporotrichosis, mucormycosis, dermatophytoses and Pneumocystis infections (all claimed). In particular, the fungal infection involves Candida, Aspergillus and/or Cryptococcus spp.; whose growth is not effectively inhibited by contact with human chitinase alone.

Sequence 1656 BP: 365 A: 497 C: 447 G: 347 T: 0 other: XX

Query Match	27.98;	Score 425.2;	DB 24;	Length 1656;
Best Local Similarity	62.08;	Pred. No. 1.1e-118;		
Matches 692; Conservative	0;	Mismatches 418;	Indels 6;	Gaps 1;
QY	23	CAGGTCTGGGCTCTTCTGCTCAATGCTCAGCTGGGGTCTGCCTACAACTCTGATATGCTATT	82	
DB	49	CAGTTTCATGGTCTGCTGCTGATGATCCATGGGGCTCTGCTGCAAACTGGTCTGCTACT	108	
QY	83	TCACCAACTGGGCCAGTATPCGGCCAGGCTCTGGGAGCTTCAAGCCCTGATGACATTAACC	142	
DB	109	TCACCAACTGGGCCAGTACAGACAGGGGAGGCTGCTCTCTGCCAAGGACTTTGGACC	168	
QY	143	CTTGCCCTGTGTACTCACTCATCTATGCTCTTCTGCTGGGATGCAGAAACAATGAGATCACC	202	
DB	169	CCAGGCTTTGCACCCCACTCATCTACGCTTCGCTGGCATGACCAACCAACGAGCTGAGCA	228	
QY	203	CCATAGAAATGGAATGATGTTACTCTCTATAAAGCTTTCAATGACTTGAATAACAGGAACA	262	
DB	229	CCACTGAGTGGATGACGAGACTCTTACAGAGTTCAATGGCTGAAGAGATGAATC	288	
QY	263	GCAAACTGAAAACCCCTCTCGCAATTTGGAGGCTGGAACTTTTGGAACTGCTCTTTTCACTA	322	
DB	289	CCAAGCTGAAGACCCCTTTAGCCATCAGGAGCTGGAGCTTGAATTTTCAGCACTCAGAAAGTTCAC	348	
QY	323	CCATGGTTTCCACTTCTCAGAACGCCAGACCTTCATTACCTCAGTTCATCAAAATTTCTGC	382	
DB	349	ATATGGTAGCCAGGCCCAACCGTCAGACCTTTGTCAACTCGGCCCTCAGGTTTCTGC	408	
QY	383	GTCAAGTATGGTTTGTAGGACTGGACCTGGAGCTGGGAATACCCAGGCTCAGCTGGGAGCC	442	
DB	409	GCAAAATACAGCTTTGACGGCTTGACCTTGACTGGGAGTACCCAGGAAGCCAGGGAGCC	468	
QY	443	CTCCTCAGGACACAGCACTCTCTTCACTGTCTGTGTAAGGAATTCGCTGAAGCTTTTGAGC	502	
DB	469	CTGCCCTAGACAGGAGCGCTTACAAACCCCTGCTACAGGACTTTGGCCAAATGCTTTCCAGC	528	
QY	503	AGGAGGCTATTGAGACCAACAGGCCCAAGCTGATGGTTACTGCTGCTGTAAGCTGGTGGGA	562	
DB	529	AGGAAGCCAGACCTCAGGGAAGAACGCCCTCTTCTGAGTGGAGCGGTTCCAGCTGGCC	588	
QY	563	TTTTCAACATCCAGGCTGGCTATGAGATCCCTGAACTTTCTAAGTACTCTGGATTTTCAATCC	622	
DB	589	AGACCTATGTGGATGCTGGATACAGGTGGACAAAATCGCCAGAACCTTGGATTTTGTCA	648	
QY	623	ATGTCATGACATATGACCTCCATGGCTCTCTGGGAGGGCTACACTGGGGAGAATAGTCCCTC	682	
DB	649	ACCTTATGGCTACGACTTCCATGGCTCTTTGGGAGAAGGTCACGGGACATAACAGCCGCC	708	
QY	683	TTTTACAAATACCTACTGACACTGGTAGCAATGCCTACCTCAATGCTGGATTTATGTCATGA	742	
DB	709	TCTACAAAGGCAAGAGAGAGTGGTGCACAGCCAGCCCTCAAGCTGGATGCTGCTGCTGC	768	
QY	743	ACTATTGGGAAGACAAATGGAGGCCCAAGCTCAGAGAGCTCATTTGTTGGATTTCCCAAGATATG	802	
DB	769	AACAGTGGCTGCAAGAGGGGACCCCTGCCAGCAAGCTGATCTCTTGGCATGCTCTACCTACG	828	

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2003, 04:38:16 ; Search time 73.5873 seconds
(without alignments)
6355.472 Million cell updates/sec

Title: US-10-004-219B-3

Perfect score: 1525

Sequence: 1 atggccaagctactctcgt.....taaaattgtagcacaaca 1525

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA.*

- 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/PCRU5_COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	426.8	28.0	1636	US-09-039-198A-1	Sequence 1, Appli
2	426.8	28.0	1636	US-08-877-599-1	Sequence 1, Appli
3	426.8	28.0	1636	US-09-267-574-1	Sequence 1, Appli
4	426.8	28.0	1643	US-08-486-839-3	Sequence 3, Appli
5	426.8	28.0	1643	US-09-151-011-3	Sequence 3, Appli
6	426.8	28.0	1643	US-08-343-623-3	Sequence 3, Appli
7	426.8	28.0	1713	US-08-486-839-5	Sequence 5, Appli
8	426.8	28.0	1713	US-09-151-011-5	Sequence 5, Appli
9	426.8	28.0	1713	US-09-343-623-5	Sequence 5, Appli
10	425.2	27.9	1656	US-09-039-198A-3	Sequence 3, Appli
11	425.2	27.9	1656	US-08-877-599-3	Sequence 3, Appli
12	425.2	27.9	1656	US-09-267-574-3	Sequence 3, Appli
13	343.2	22.5	1433	US-08-694-915-1	Sequence 1, Appli
14	330.8	21.7	1526	US-08-694-915-3	Sequence 3, Appli
15	314.8	20.6	1681	PCT-US94-07754-4	Sequence 4, Appli
16	311.6	20.4	1681	US-08-581-527-4	Sequence 4, Appli
17	140	9.2	2452	US-08-524-051-1	Sequence 1, Appli
18	126	8.3	1677	US-09-545-814-13	Sequence 13, Appli
c 19	126	8.3	1677	US-09-545-814-15	Sequence 15, Appli
c 20	126	8.3	1749	US-09-545-814-4	Sequence 4, Appli
c 21	126	8.3	1749	US-09-545-814-6	Sequence 6, Appli
c 22	126	8.3	2610	US-09-545-814-1	Sequence 1, Appli
c 23	126	8.3	2610	US-09-545-814-3	Sequence 3, Appli
c 24	124.4	8.2	1919	US-09-545-814-31	Sequence 31, Appli
c 25	124.4	8.2	1919	US-09-545-814-33	Sequence 33, Appli
c 26	121.4	8.0	1608	US-09-292-225-20	Sequence 20, Appli
c 27	121.4	8.0	1608	US-09-292-225-22	Sequence 22, Appli

28	121.4	8.0	1665	4	US-09-292-225-17	Sequence 17, Appli
c 29	121.4	8.0	1665	4	US-09-292-225-19	Sequence 19, Appli
c 30	121.4	8.0	1752	4	US-09-292-225-14	Sequence 14, Appli
c 31	121.4	8.0	1752	4	US-09-292-225-16	Sequence 16, Appli
c 32	112.8	7.4	1470	4	US-09-292-225-40	Sequence 40, Appli
c 33	112.8	7.4	1470	4	US-09-292-225-42	Sequence 42, Appli
c 34	112.8	7.4	1527	4	US-09-292-225-37	Sequence 37, Appli
c 35	112.8	7.4	1527	4	US-09-292-225-39	Sequence 39, Appli
c 36	112.8	7.4	1621	4	US-09-292-225-34	Sequence 34, Appli
c 37	112.8	7.4	1621	4	US-09-292-225-36	Sequence 36, Appli
c 38	109.6	7.2	1478	4	US-09-545-814-28	Sequence 28, Appli
c 39	109.6	7.2	1478	4	US-09-545-814-30	Sequence 30, Appli
40	64	4.2	147	4	US-09-156-856-8	Sequence 8, Appli
41	54.8	3.6	1167	1	US-07-939-501A-6	Sequence 6, Appli
42	54.8	3.6	1167	4	US-08-448-398-10	Sequence 10, Appli
43	54.8	3.6	1320	1	US-07-939-501A-14	Sequence 14, Appli
44	54.8	3.6	1364	1	US-07-939-501A-13	Sequence 13, Appli
45	54.8	3.6	1405	1	US-07-939-501A-9	Sequence 9, Appli

ALIGNMENTS

RESULT 1
US-09-039-198A-1
; Sequence 1, Application US/09039198A
; Patent No. 6200951
; GENERAL INFORMATION:
; APPLICANT: Gray, Patrick W.
; APPLICANT: Tjoelker, Larry W.
; TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive/6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/039,198A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 27866/34391
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1636 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..1399
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 65..1399
US-09-039-198A-1

Query Match 28.0%; Score 426.8; DB 4;
Best Local Similarity 62.1%; Pred. No. 6.6e-123;
Matches 693; Conservative 0; Mismatches 417; Indels 6; Gaps 1;

Db 204 CCAGTGAAGTGAATGACGAGACTCTTACCAGGAGTTCAATGGCTTGAAGAAGATGAATC 263
QY 263 GCAAACTGAAACCCCTCTTGGCAATTTGGAGGCTGGAACTTTGGAACCTGCTCTTTCACATA 322
Db 264 CCAAGCTGAAGACCCCTTGTAGCCATCGGAGGCTGGAAATTTGGCACTCAGAGTTCCACAG 323
QY 323 CCATGGTTTCCACTTCTCAGAACCCGACAGCCCTTCATTACCTCAGTCATCAAAATTTCTGC 382
Db 324 ATATGGTAGCCACGGCCCAACACCGTTCAGACCTTTGTCAACTCGGCCATCAGTTTCTGC 383
QY 383 GTCAAGTATGGTTTGTATGACTGGACCTGGACCTGGGAATACCCAGGCTCAGCTGGAGCC 442
Db 384 GCAAAATACAGCTTTGACGGCCCTTGACCTTGGAGTACCCAGGAGCCAGGGGAGCC 443
QY 443 CTCCTCAGGACAGCATCTCTTCACTGCTCTGTGTGAAGGAATCGGTGAAGCTTTTGAGC 502
Db 444 CTGCGGTAGCAAGGAGCCCTTCACAAACCCTGGTACAGGCTTGGCCAATGCTTCCAGC 503
QY 503 AGGAGCTATTGAGACAAACAGGCCAGACTGATGTTACTGCTGTAGCTGTGGTGGGA 562
Db 504 AGGAAGCCAGACCTCAGGGAAGGAGCGCTTCTTCTGAGTGCAGCGGTTCCAGCTGGGC 563
QY 563 TTTCCAACTCAGGCTGCTATGAGATCCCTGAACTTTCTAGTACCTGGATTTTCATCC 622
Db 564 AGACCTATGTGATGCTGATACGAGGTGGACAAATCGCCAGAACCTTGGATTTGTGA 623
QY 623 ATGTATGATATGATACCTCCATGGCTCTGGAGGCTTACACTGGGGAGAAATAGTCCTC 682
Db 624 ACCTTATGGCTACGACTTCCATGGCTTTGGGAGAGGTACGGGACATACAGGCCCC 683
QY 683 TTTTCAAAATPACCTACTGAGACTGAGCAATGCCCTACCTCAATGTGGATTTATGTCATGA 742
Db 684 TCTACAAGAGGAAGAGAGTGGTGCAGCAGCCAGCCCTCAACGTGGATGCTGTGTGC 743
QY 743 ACTATTGGAGAACATGAGCCCGCCAGCTGAGAGCTCATTTGGATTCCAGAGATG 802
Db 744 AACAGTGGTGCAGAGGGGAGCCCTGCGCAGCAAGCTGATCTTGGCATGACCTACATCG 803
QY 803 GACACACCTTCTCTGAGAAACCCCTCTGATAATGGAATTTGGTGGCCCTACCTCTGTGG 862
Db 804 GAGCTCTCTTCACTTGGCTCTCATCAGACACAGAGTGGGGGCCCCAGCCAGAGGT 863
QY 863 ATGGCCCTGCTGGCCCTATACAGACAGGCTGGGTTCCTGGCCCTACTATGAGATTGCA 922
Db 864 CTGGCACTCCAGGCCCTTCCACCAAGGAAGAGGATGCTGGCCCTACTATGAAGTCTGCT 923
QY 923 CTTTCTGAGAGTGGAGCCACTGAGGTCTGGATGCCCTCCCAAGAGTGGCCCTATGCT 982
Db 924 CCT-----GGAGGGGGCCCAACAGAGAAATCCAGGATCAGAGAGTGGCCCTACATCT 977
QY 983 ATAGGCCAACAGTGGCTTGGCTATGACAATATCAAGAGCTTTCAGTGTAAAGGCTCAGT 1042
Db 978 TCCGGACAACAGTGGGTGGCTTGTATGATGATGAGAGCTTCAAAACCAAGGTCAGCT 1037
QY 1043 GGTAAAGCAGAACAAATTTGGAGTGCATGATCTGGCCCATTTACCTTGTATGATTTCA 1102
Db 1038 ATCTGAAGCAGAGGAGTGGCGGGCCATGGCTGGCACTGGACTTAGATGACTTTG 1097
QY 1103 CTGGCTCTTCTGTATGATAGGGAATTTTCCTCTGA 1138
Db 1098 CCGGCTTCTCTGCAACACGAGGGCCGATACCCCTCA 1133

RESULT 3

US-09-267-574-1
; Sequence 1, Application US/09267574
; Patent No. 639571
; GENERAL INFORMATION:
; APPLICANT: Gray, Patrick W.
; APPLICANT: Tjoelker, Larry W.
; TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
; FILE REFERENCE: 27866/35407
; CURRENT APPLICATION NUMBER: US/09/267,574

; CURRENT FILING DATE: 1999-03-12
; EARLIER APPLICATION NUMBER: 09/039,198
; EARLIER FILING DATE: 1998-03-12
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1636
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(1399)
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (65)..(1399)
US-09-267-574-1

Query Match 28.0%; Score 426.8; DB 4; Length 1636;
Best Local Similarity 62.1%; Pred. No. 6.6e-123;
Matches 693; Conservative 0; Mismatches 417; Indels 6; Gaps 1;

QY 23 CAGGCTGCTCTTCTGCTGAATGCTCAGCTGGGCTCTGCCTACATCTGATATGCTATT 82
Db 24 CAGGTTTCATGCTCTGCTGCTGATGCCATGGGCTCTGCTGCANAACTGGTCTGCTACT 83
QY 83 TCACCAACTGGGCCCAAGTATCGGCCAGGTCTGGGAGCTTCAAGCCTGATGACATTAACC 142
Db 84 TCACCAACTGGGCCCAAGTACAGACAGGGGAGGCTCGCTCTGCCCAAGGACTTGGACC 143
QY 143 CTGCTGCTGTACTACCTGATCTATGCTTTGCTGGGATGAGCAACAATGAGATCACCA 202
Db 144 CCAGCTTTTGCAACCCACCTCATCTAGCCCTTCTGCTGGCATGACCAACACCCAGCTGAGCA 203
QY 203 CCATAGAATGGATGATGTTACTCTCTATAAAGCTTTCAATGATCTTGAAGAAACAGGACA 262
Db 204 CCACGTAGGGAATGACGAGACTCTCTACAGAGTTCATAGGCTTGAAGAAGATGAATC 263
QY 263 GCAAACTGAAACCCCTCTTGGAGGCTTGGAACTTTTGAACCTGCTCTTTCACATA 322
Db 264 CCAAGCTGAAGACCTTGTAGCCATCGGAGCTTGGAAATTCGSCACTCAGAACTTCACAG 323
QY 323 CCATGGTTTCCACTTCTCAGAACCCGACAGCTTCAATACCTCAGTCATCAAAATTTCTGC 382
Db 324 ATATGTTAGCCAGGCCCAACACCGTTCAGACCTTTTGTCAACTCGGCCATCAGGTTTCTGC 383
QY 383 GTCAAGTATGGTTTGTAGTGGACTGGACCTGGAGTGGGAATACCCAGGCTCAGCTGGGAGCC 442
Db 384 GCAAAATACAGCTTTGACGGCTTGAACCTTGAAGTGGAGTACCCAGGAAGCCAGGGAGCC 443
QY 443 CTCTCTCAGGACAGCATCTCTTCACTGTCTGCTGGAAGGAATTCGCTGAAGCTTTTGGAGC 502
Db 444 CTGCGCTTACCAAGGAGCCCTTCAACACCCCTGTCAGGACTTGGCCCATGCTCTCCAGC 503
QY 503 AGGAGCTATTGAGACAAACAGGCCAGACTGATGTTTACTGCTGTAGTGTGGTGGGA 562
Db 504 AGGAAGCCAGACCTCAGGGAAGGAACGCTTCTTCTGAGTGCAGCGGTTCACAGCTGGGC 563
QY 563 TTTTCAACATCAGGCTGCTGATGATCCCTGAACCTTTCTAGTACCTGGATTTTCATCC 622
Db 564 AGACCTATGTGGATGCTGGATAGGAGTGGACAAATTCGCCCAAGAACCTTGGATTTGTCA 623
QY 623 ATGCTCATGATATGATACCTCCATGGCTCTGCGGAGGCTTACACTGGGGAGAAATAGTCCTC 682
Db 624 ACCTTATGGCTACGACTTCCATGGCTTGGGAGAGGTCAGGGGACATACAGAGCCCC 683
QY 683 TTTTCAAAATPACCTACTGAGACTGGTAGCAATGCCCTACCTCAATGTGGATTTATGTCATGA 742
Db 684 TCTACAAGAGGAAGAGAGTGGTGCAGACCCAGCCAGCTTCAACGTGGATGCTGTGTGC 743
QY 743 ACTATTGGAGAACAAATGAGCCCGCCAGCTCAGAGCTCATTTGGATTTCCCAAGATG 802
Db 744 AACAGTGGTGCAGAGGGGAGCCCTGCCAGCAAGCTGATCTTGGCATGCTCTTACCTAC 803

QY 803 GACACACCTTCATCTCGAGAACCCCTCTGATGATGAATGGTGGCCCTACCTCTGGTG 862
DB 804 GACGCTCTTCACACTGGCTCTCATCAGACACAGAGTGGGGCCCCAGCCACAGGCT 863
QY 863 ATGGCCCTGTGGGCTTATACAGACAGGCTGGTCTGGGCTTACTATGAGATTGCA 922
DB 864 CTGGCACTCCAGGCCCTTCCACAAAGGAAGGAGGATGCTGGGCTTACTATGAAGTCTGCT 923
QY 923 CCTTTCTGAGAGTGGAGCCACTGAGTCTGGGATGCTCCCAAGAGTGGCCCTATGCCT 982
DB 924 CCT-----GGAAGGGGGCCACCAACAGAGATCCAGGATCAGAGGTGCCCTACATCT 977
QY 983 ATAAGGCCAACAGTGGCTTGGCTATGACAAATACAGAGCTTCAGTGTAAAGGCTCAGT 1042
DB 978 TCCGGGACAAACAGTGGTGGCTTGTATGATGTGGAGAGCTTCAAAAACCAAGTCACT 1037
QY 1043 GGCTTAAGCAGAACATTTTGGAGTGGCCATGATCTGGGCCATGACCTTGTAGTCA 1102
DB 1038 ATCTGAAGCAGAGGAGGACTGGGGGGCCATGCTGGGCACTGGACTTAGATGACTTTG 1097
QY 1103 CTGGCTCTTCTGTGATCAGGGAATAATTTCTCTGA 1138
DB 1098 CCGGCTCTCTGCACACGAGCGGCGATACCCCTCA 1133

RESULT 4

US-08-486-839-3

; Sequence 3, Application US/08486839

; Patent No. 5928928

; GENERAL INFORMATION:

; APPLICANT:

; TITLE OF INVENTION: A human chitinase, its recombinant

; TITLE OF INVENTION: production, its use for decomposing chitin, its use

; TITLE OF INVENTION: in therapy or prophylaxis against infection diseases.

; NUMBER OF SEQUENCES: 16

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Hoffmann & Baron

; STREET: 350 Jericho Turnpike

; CITY: Jericho

; STATE: New York

; COUNTRY: United States of America

; ZIP: 11758

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/486,839

; FILING DATE: 07 - June - 1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Baron, Ronald J.

; REGISTRATION NUMBER: 29,281

; REFERENCE/DOCKET NUMBER: 294-26

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (516) 822-3550

; TELEFAX: (516) 822-3582

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1643 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: unknown

; TOPOLOGY: unknown

; MOLECULE TYPE: cDNA

; HYPOTHETICAL: NO

US-08-486-839-3

Query Match

Best Local Similarity 28.0%; Score 426.8; DB 2; Length 1643;

Matches 693; Conservative 0; Mismatches 417; Indels 6; Gaps 1;

QY 23 CAGGTCTGCTCTTCTGCTGAATGCTCAGCTGGGGTCTGCCTACAAATCTGATATGCTATT 82

DB 35 CAGGTCTGCTCTGCTGATGATCCATGGGCTCTGCTCCAAACTGGTCTGCTACT 94
QY 83 TCACAACTGGGCGAGTATCGGCGAGGCTTCAGAGCTTCAGAGCTGATGACATTAAC 142
DB 95 TCACAACTGGGCGAGTACAGACAGGGGAGGCTCGCTTCTTCCCTCCCAAGAGCTTGAC 154
QY 143 CTTGCTGTCTACTCACTGATCTATGCTCTTCTGCGATGAGAGCAAGTATGATACCA 202
DB 155 CAGGCTTTCACCCACTCATCTACGCTTTCGCTGCTGATGACCAACACCACTGAGCA 214
QY 203 CCATAGAAATGATGTTTACTCTCTATATAAGCTTTCAATGACTTTGAAACACAGGA 262
DB 215 CCAGTGTGAAATGACGAGACTCTCTACAGGAGTTCATGCTTGAAGAGATGAATC 274
QY 263 GCAAACTGAAACCCCTCTGGCAATTTGGAGGCTTGAACITTTGAACTGCTCTTCACTA 322
DB 275 CCAAGCTGAAGACCTTTAGCCATCGGAGGCTTGAATTTTCGGCACTCAGAAGTTCACAG 334
QY 323 CCATGTTTCCACTTCTCAGAACCCGACAGCTTCATTAACCTCAGTCATCAAAATTTCTGC 382
DB 335 ATATGTAGCCAGCGCCACACACCGTCAGACCTTTGTCACTCGGCCATCAGGTTTCTGC 394
QY 383 GTCAGTATGGTGTGATGACTGGACCTGGACTTGGAAATACCAAGGCTTCAGTGGAGCC 442
DB 395 GCAAAATACAGCTTTGACGGCTTTGACCTTGACTGGAGTACCAGGAAGCCAGGGAGCC 454
QY 443 CTCTCAGGACAGCACTCTCTCAGTCTCTGCTGAGGAAATGCTGAGCTTTTTCAGC 502
DB 455 CTGCTGACAAAGGAGGCTTTCACAACTGTTAGGACTTTGGCAATGCTTTCAGC 514
QY 503 AGGAGGCTATTGAGACACAGCGCCAGACTGATGTTTACTGCTGTAGCTGGTGGGA 562
DB 515 AGAAGCCAGACTCAGGGAAGAGCGCTTCTTCTGAGTGGAGGCTTCAGAGTGGGC 574
QY 563 TTTTCAACATPCAGGTGCTATGAGATCCCTGAACTTTCTAAGTACCTGGATTTCATCC 622
DB 575 AGACCTATGTGGATGCTGGATAGAGTGGACAAAATCGCCAGAACCTTGGATTGTCA 634
QY 623 ATGTCATGACATGACCTCCATGCTCTCGGAGGCTACACTGGGAGAAATAGTCTC 682
DB 635 ACCTTATGGCTACGACTTCCATGGCTCTTGGGAGAGGTACCGGACATACAGCCCC 694
QY 683 TTTTCAAAATACCTACTGAGACTGGTAGCAATCCCTACCTCAATATGGATTTATGTCATGA 742
DB 695 TCTACAAGAGGCAAGAGAGAGTGGTGCAGACAGCCCTCAACGTGGATGCTGCTGTC 754
QY 743 ACTATTTGGAGAACATGAGGCCCCAGCTGAGAGCTCATTTGTTGATTTCCAGAGATG 802
DB 755 AACAGTGGCTGACAGAGGGGACCCCTGCCAGCAAGCTGATCTTGGCATGCCCTACCTAG 814
QY 803 GACACACCTTCTCTGAGAAACCCCTCTGATGATGAAATGGTGGCCCTACCTCTGGTG 862
DB 815 GAGCTCTCTACACTGGCTCTCTATCAGACACAGAGTGGGGGCCCCAGCACAGGCT 874
QY 863 ATGGCCCTGCTGGGCTTATACAGACAGGCTGGGTTCTGGGCTTACTATGAGATTGCA 922
DB 875 CTGGCACTCCAGGCCCTTTCACCAAGGAAGGAGGATGCTGGGCTTACTATGAAGTCTGCT 934
QY 923 CTTTCTGAGAGTGGAGCCACTGAGTCTGGGATGCCCTCCCAAGAGTGGCCCTATGCTCT 982
DB 935 CCT-----GGAAGGGGGCCACCAACAGAGATCCAGGATCAGAGGTGCCCTACATCT 988
QY 983 ATAAGGCCCAAGTGGCTTGGCTATGACAAATATCAAGAGCTTTCAGTGTAAAGGCTCAGT 1042
DB 989 TCGGGACAAACAGTGGGCTTGTGATGATGTGGAGAGCTTCAAAACCAAGGTCAGCT 1048
QY 1043 GCTTAAAGCAGAACAAATTTTGGAGTGGCCATGATCTGGGCCATTTGACCTTGTACTTCA 1102
DB 1049 ATCTGAAGCAGAAAGGAGTGGGGGGCCATGCTCTGGGCACTGGGACTTAGATGACTTTG 1108
QY 1103 CTGGCTCTTCTGATGATCAGGGAATAATTTCTCTGA 1138

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

APPLICATION DATA: US/09/343, 623
FILING DATE: 07-June-1995

APPLICATION NUMBER: US/08/486, 839
FILING DATE: 07-June-1995

ATTORNEY/AGENT INFORMATION:
NAME: Baron, Ronald J.

REGISTRATION NUMBER: 29,281
REFERENCE/DOCKET NUMBER: 294-26

TELEPHONE: (516) 822-3550
TELEFAX: (516) 822-3582

INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:

LENGTH: 1643 base pairs
TYPE: nucleic acid

STRANDEDNESS: unknown
TOPOLOGY: unknown

MOLECULE TYPE: CDNA
HYPOTHETICAL: NO

US-09-343-623-3

Query Match 28.0%; Score 426.8; DB 4; Length 1643;

Best Local Similarity 62.1%; Pred. No. 6.6e-123;

Mismatches 693; Conservative 0; Gaps 6; Indels 1;

QY 23 CAGGTCTGGCTCTCTCTGCTGAATGCTCAGCTGGGGTCTGCTACAATCTGATGCTATT 82
DB 35 CAGGTCTGCTCTCTGCTGATGCTCCATGGGCTCTGCTCCAAACTGCTGCTACT 94
QY 83 TCACCACTGGGCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 142
DB 95 TCACCACTGGGCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 154
QY 143 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 202
DB 155 CAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 214
QY 203 CCATAGATGGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 262
DB 215 CCATGATGGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 274
QY 263 GCAACTGGAACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 322
DB 275 CCAAGCTGAAGACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 334
QY 323 CCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 382
DB 335 ATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 394
QY 383 GTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 442
DB 395 GCAATACAGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 454
QY 443 CTGCT 502
DB 455 CTGCT 514
QY 503 AGGAGCTATTGAGACCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 562
DB 515 AGGAGCTATTGAGACCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 574
QY 563 TTTCACATCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 622
DB 575 AGACCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 634
QY 623 ATGCTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 682

DB 635 ACCTATGGCTACGACTTCCATGGCTTTGGGAGAGGTCACGGGACATACAGGCCCC 694
QY 683 TTTCACAAATACCTACTGAGACTGGTAGCAATGCCTACCTCAATGTGATGATGCTCATGA 742
DB 695 TCACAAAGAGGCAAGAGAGAGTGGTGACAGCCGCTCAAGTGGATGCTGCTGCTGC 754
QY 743 ACTATTGGAGAAATGGAGCCGCTGAGAGCTCAATGTTGGATTCCCAAGAGATG 802
DB 755 AACAGTGGCTGACAGAGGGGACCCCTGCCAGCAAGCTGATCTTTGGCATGCTACCTACG 814
QY 803 GACACACTTCATCTGAGAAACCCCTCTGATATGGAATGTTGGTCCCTACCTCTGCTG 862
DB 815 GACGCTCTTACACTGGCTCTCATCAGACACAGAGTGGGGGCCCCACACAGGCT 874
QY 863 ATGGCCTGCTGGCGCTATACACAGACAGGCTGGGTTCTGGGCTTACTATGAGATTGCA 922
DB 875 CTGGCACTCCAGGCCCTTCCACCAAGGAGGAGTGGTGGCTTACTATGAGTCTGCT 934
QY 923 CTTTCTGAGAAGTGGAGCCACTGAGGTCTGGGATGCTCCCAAGAGTCCCTATGCT 982
DB 935 CCT-----GGAAGGGGGCCACCAACAGAGATCCAGGATCAGAAGTCCCTACATCT 988
QY 983 ATAGGCCAACGAGTGGCTTGGCTATGACAAATATCAAGACTTCAGTTTAAAGCTCAGT 1042
DB 989 TCCGGGACCAACAGTGGTGGGCTTTGATGATGTTGAGAGCTTCAAAACCAAGGTCAGT 1048
QY 1043 GGCTTAAGCAGAAATTTTGGAGTGCCATGATCTGGGCAATGACCTTGTGATGACTTCA 1102
DB 1049 ATCTGAGCAGAGGACTGGGGGGCCATGCTGCTGGGCACTGGGACTTGTGATGACTTTG 1108
QY 1103 CTGGCTCTTCTGATGACGAGAAATTTCTCTCTGA 1138
DB 1109 CCGGCTTCTCTCAACAGGGGCGGATACCCCTCA 1144

RESULT 7

US-08-486-839-5

Sequence 5, Application US/08486839

Patent No. 5928928

GENERAL INFORMATION:

APPLICANT:

TITLE OF INVENTION: A human chitinase, its recombinant

TITLE OF INVENTION: production, its use for decomposing chitin, its use

TITLE OF INVENTION: in therapy or prophylaxis against infection diseases.

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESSEE: Hoffmann & Baron

STREET: 350 Jericho Turnpike

CITY: Jericho

STATE: New York

COUNTRY: United States of America

ZIP: 11758

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/486,839

FILING DATE: 07 - June - 1995

ATTORNEY/AGENT INFORMATION:

NAME: Baron, Ronald J.

REGISTRATION NUMBER: 29,281

REFERENCE/DOCKET NUMBER: 294-26

TELEPHONE: (516) 822-3550

TELEFAX: (516) 822-3582

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 1713 base pairs

TYPE: nucleic acid

STRANDEDNESS: unknown

Query Match	28.0%	Score 426.8;	DB 3;	Length 1713;
Best Local Similarity	62.1%	Pred. No. 6.8e-123;		
Matches 693; Conservative	0;	Mismatches 417;	Indels 6;	Gaps 1;
QY	23	CAGGTCTGGCTCTTCGCTGAATGCTCAGCTGGGGTCTGCCTACAATCTGATATGCTACT	82	
Db	35	CAGTTTCATGTCCTGCTGATGATCCATGGGGCTCTGCTCCAAACTGGCTGCTACT	94	
QY	83	TCACCAACTGGGCCCCAGTATCGGCCAGGTCCTGGGAGCTTCAAGCCTGTGATGACATTAACC	142	
Db	95	TCACCAACTGGGCCCCAGTACAGACAGGGGAGGCTCGCTTCTGCCCAAGGACTTTGGAAC	154	
QY	143	CCTGCCTGTGCTACTCACCTGATCTATGCCCTTTTCTGGGATGCGAACAATGAGATCACCA	202	
Db	155	CCAGCCTTTGCACCCACCTCATCTAGCCCTTCGCTGGCATGACCAACCCACCGCTTGAGCA	214	
QY	203	CCATAGAATGGAAATGATGTTACTCTCTATAAAGCTTTTCAATGACTTCGAAAAACAGGAACA	262	

Db 215 CCCTGAGTGAATGACGAGACTCTACAGGAGTTCAATGGCTGAAGAAGATGATC 274
Qy 263 GCAAACTGAAACCTCTCTGCAATTTGGAGCTGAATTTGGAGCTGCTCTTCACTA 322
Db 275 CCAAGCTGAGACCTCTTAGCCATCGAGGCTGGAATTTGGGCACCTCAAGTCTACAG 334
Qy 323 CCATGGTTCCACTTCTCAGAACCGCCAGACTTTCATTACCTCAGTCAATCAAAATTTCTGC 382
Db 335 ATATGGTAGCCACGCGCAACACCGTCAGACCTTTGTCACTCGGCCATCAGTTTCTGC 394
Qy 383 GTAGATGTTGATGGAGTGGACCTGGACCTGGGGAATACCCAGGCTCAGTGGGAGCC 442
Db 395 GCAAAATACAGCTTTGACGGCTTACCTTGACTGGGAGTACCCAGGAAGCCAGGGAGCC 454
Qy 443 CTCCTCAGGACAGCTCTCTCAGTCTCTGTTGGAAGAAATCGTGAAGCTTTTCAGC 502
Db 455 CTGCGGTAGACAGGAGCGCTTACAAACCTGGTACAGGACTTGGCCATGCGCTTCAGC 514
Qy 503 AGGAGGCTATGAGAGCAACAGGCCACAGACTGATGGTTACTGTCTGTAGCTGGTGGGA 562
Db 515 AGGAAGCCAGACCTCAGGGAAGGAGCGCTTCTCTGAGTGCAGCGGTTCCAGCTGGC 574
Qy 563 TTTCACATCCAGGCTGGCTATGAGATCCCTGAATTTCTAAGTACTGATGATTCATCC 622
Db 575 AGACCTATGAGGATGCTGATGAGAGGTGGACAAATCGGCCAGAACCTGGATTTGCA 634
Qy 623 ATGTCATGACATGACCTCCAGTCCCTGCTGGAGGCTACACTGGGGAAGTATGCTC 682
Db 635 ACCTATGGCTAGGACTTCCATGGCTTGGGAGGAGTGCAGGGACATTAAGAGCCCC 694
Qy 683 TTTACAAATACCTACTGAGAGTGGTGAATGCTACCTCAATGTGGATTATGTCATGA 742
Db 695 TCTACAGAGGCAAGAGAGAGTGGTGCAGACGCCCTCAAGTGGATGCTGCTGC 754
Qy 743 ACTATTGGAAGACATGAGAGCCAGCTGAGAGCTCAATGTTGGATTCCAGAGTATG 802
Db 755 AACAGTGGCTGCAAGAGGAGGACCCCTGCCAGCAAGCTGATCCTTGGCATGCCCTAC 814
Qy 803 GACACACTTTCATCTGAGAAACCCCTCTGATATGGAATTTGGTCCCTACCTCTGGT 862
Db 815 GACGCTCTTCACTGAGCTGCTCTCATCAGACACAGAGTGGGGCCCGCCAGCAGG 874
Qy 863 ATGGCCCTGCTGGCCCTATFACAGACAGGCTGGGTTCTGGGCTACTATFAGATTGCA 922
Db 875 CTGGCACTCCAGGCCCTTCCACCAAGGAAGGAGTGTGGCTACTATGAGTCTGCT 934
Qy 923 CTTTCTGAGAGTGGAGCCACTGAGGTCTGGATGCTCCCAAGAGTCCCTATGCT 982
Db 935 CCT-----GGAAGGGGGCCACCAACAGAGATCCAGGATCAGAGGTGCCCTACATCT 988
Qy 983 ATAAGGCCAAGAGTGGCTTGGCTATGACAAATATCAAGAGCTTTCAGTGTAAAGCTCAGT 1042
Db 989 TCCGGGACACACAGTGGTGGCTTTGATGATGTGGAGAGCTTCAAAACCAAGTCAAGT 1048
Qy 1043 GCGTTAAGCAGAACATTTTGGAGGTGCCATGATCTGGGCCATFAGCTTGAAGTCA 1102
Db 1049 ATCTGAAGCAGAGGAGTGGGGGGCCATGCTGGGCACTGGACTTAGATGACTTTG 1108
Qy 1103 CTGGCTCTTCTGATCAGGGAATTTCTCTGA 1138
Db 1109 CCGGCTCTCTCTGCAACAGGCGCGATACCCCTCA 1144

RESULT 9

US-09-343-623-5

; Sequence 5, Application US/09343623

; Patent No. 6303118

; GENERAL INFORMATION:

; APPLICANT:

; TITLE OF INVENTION: A human chitinase, its recombinant

; TITLE OF INVENTION: production, its use for decomposing chitin, its use

; TITLE OF INVENTION: in therapy or prophylaxis against infection diseases.

NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hoffmann & Baron
STREET: 350 Jericho Turnpike
CITY: Jericho
STATE: New York
COUNTRY: United States of America
ZIP: 11758
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
SOFTWARE:
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/343,623
FILING DATE:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US/08/486,839
FILING DATE: 07-June-1995
ATTORNEY/AGENT INFORMATION:
NAME: Baron, Ronald J.
REGISTRATION NUMBER: 29,281
REFERENCE/DOCKET NUMBER: 294-26
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 822-3550
TELEFAX: (516) 822-3582
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1713 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
US-09-343-623-5

Query Match 28.0%; Score 426.8; DB 4; Length 1713;

Best Local Similarity 62.1%; Pred. No. 6.8e-123;

Matches 693; Conservative 0; Mismatches 417; Indels 6; Gaps 1;

Qy 23 CAGGTCTGGCTCTTCTGCTGAATGCTCAGCTGGGGTCTGCTCAATCTGATATGCTATT 82
Db 35 CAGGTTTCATGCTGCTGCTGATGATCCCATGGGCTCTGCTCCAAAATGCTGCTACT 94
Qy 83 TCACCAACTGGGCCAGTATCGCCAGGCTCGGGAGCTTCAGCCCTGATGACATTAACC 142
Db 95 TCACCAACTGGGCCAGTATCAGACAGGGGAGGCTCGCTTCCTGCCCAAGGACTTGACC 154
Qy 143 CTGCTCTGTCTACTACCTGATCTATGCTTTCCTGGATGCAAGAACATGATCAGCA 202
Db 155 CCAGCCTTTGACCCACCTCATCTAGCCTTCGCTGGCATGACCAACCACCTGAGCA 214
Qy 203 CCATAGAATGAATGATGTTACTCTCTATAAGCTTTCAATGACTTTGAATAACAGGAACA 262
Db 215 CCATCAGTGAATGACGAGACTCTCTACAGGAGTTCAATGGCTTGAAGAAGATGAATC 274
Qy 263 GCAAACTGAAACCTCTCTGGCAATTTGGAGGTGGAACTTTGAACTGCTCTTCACTA 322
Db 275 CCAGCTGAAGACCTCTTAGCCATCGAGGCTGGAATTTCCGCACTCAGAACTTCACAG 334
Qy 323 CCATGTTTTCCTCTCTCAGAACCCGACAGCTTTCATTTACCTCAGTCAATCAATTTCTGC 382
Db 335 ATATGGTAGCCACGCGCAACACCGTCAGACCTTTTGTCACTGGGCCATCAGTTTCTGC 394
Qy 383 GTCAATGAGGTTTGAATGAGTGGAGCTGGAGTGGGAATACCCAGGCTCAGCTGGGAGCC 442
Db 395 GCAAAATACAGCTTTGACGGCTTGAACCTTGACTGGGAGTACCCAGGAAGCCAGGGAGCC 454
Qy 443 CTCCTCAGGACAGCTCTTTCCTGCTGCTGTTGAGGAAGTATGCTGAGCTTTTGAGC 502
Db 455 CTGCGGTAGACAGAGGAGGCTTTCACAAACCTTGTACAGGACTTGGCCAATGCTTCCAGC 514

Db 769 AACAGTGGCTGCAGAGGGGACCCCTGCCAGCAAGCTGATCTTTGGCATGCCCTACCTACG 828
QY 803 GACACACCTTTCATCTCTGAGAAACCCCTCTGATATGTAATGTAATGTCCTACCTCTGGTG 862
Db 829 GAGCTCTCTTCACTGCTGCTCTCTATCAGACACAGAGTGGGGCCCGCCAGCAGAGGT 888
QY 863 ATGGCCCTGCTGGCGCTATPACACAGACAGCTGGGTCTGGGCTCTACTATGAGATTGCA 922
Db 889 CTGGCACTCCAGGCCCTTCCACAGAGGAGGATGCTGGCTCTACTATGAGTCTGCT 948
QY 923 CTTTCTGAGAACTGAGGACCTAGAGTCTGGGATGCTCCCAAGAGTGGCCCTATGCT 982
Db 949 CTT-----GGAAGGGGGCCACCAACAGAGATCCAGGATCAGAGGTGGCCCTACATCT 1002
QY 983 ATAAAGCCACAGTGGCTGCTATGATCAATATCAAGAGCTTCAAGTGTAAAGCTCAGT 1042
Db 1003 TCCGGACACACAGTGGGTGGTGTGATGATGGAGAGCTTCAAAACCAAGGTGAGCT 1062
QY 1043 GGCTTAAGCAGAACAAATTTGGAGGTGCCATGATCTGGGCCATTGACCTTGTGACTTCA 1102
Db 1063 ATCTGAAGCAGAGGAGCTGGCGGGCCATGCTGGGCACTGGACTTAGATGACTTTG 1122
QY 1103 CTGGCTCTTCTGATCAGGGAATTTCTCTCTGA 1138
Db 1123 CCGGCTTCTCTCAACACAGGCGCCATACCCCTCA 1158

RESULT 11
US-08-877-599-3
; Sequence 3, Application US/08877599
; Patent No. 6372212
; GENERAL INFORMATION:
; APPLICANT: Gray, Patrick W.
; TITLE OF INVENTION: Chitinase Materials and Methods
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: US/08/877,599
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/663,618
; FILING DATE: 14-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 27866/33994
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1656 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 27..1424
; FEATURE:

; NAME/KEY: mat_peptide
; LOCATION: 90..1424
US-08-877-599-3
Query Match 27.9%; Score 425.2; DB 4; Length 1656;
Best Local Similarity 62.0%; Pred. No. 2.1e-122;
Matches 692; Conservative 0; Mismatches 418; Indels 6; Gaps 1;
QY 23 CAGGTCTGGCTCTTCTGCTGAATGCTCAGCTGGGTCTGCTACAACTCTGATGCTATT 82
Db 49 CAGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 108
QY 83 TCACCAACTGGGCCAGTATCGGCCAGGTCTGGGGAGCTTCAAGCCTGATGACATTAAAC 142
Db 109 TCACCAACTGGGCCAGTATCGGCCAGGTCTGGGGAGGTCTGCTTCTGCTGCTGCTGCT 168
QY 143 CTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 202
Db 169 CCAGCCTTTTGCACCCACCTCATCTAGCCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 228
QY 203 CCATAGAACTGAATGATGCTTACTCTCTATAAAGCTTCAATGACTTGAAGAAACAGGA 262
Db 229 CCACTGAGTGAATGAGAGACTCTCTACCAAGAGTTCATGCTGCTGCTGCTGCTGCTGCT 288
QY 263 GCAAACTGAAACCCCTCTGGCAATTTGGAGCTTGGAACTTCTGCTGCTGCTGCTGCTGCT 322
Db 289 CCAAGCTGAAGACCTTGTAGCCATCGGAGGTGGAATTTTCAAGCTCAGAAAGTTCACAG 348
QY 323 CCATGCTTCCACTTCTCAGAACCCGACAGCTTCTTCACTGCTGCTGCTGCTGCTGCTGCT 382
Db 349 ATATGCTAGCCAGCCCAACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 408
QY 383 GTCAGTATGGTGTGATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 442
Db 409 GCAATACAGCTTGTGAGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 468
QY 443 CTCTCAGGACAAAGCATCTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 502
Db 469 CTGCGCTAGACAAAGGAGCTTTCACAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 528
QY 503 AGGAGCTATTGAGACACAGCCGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 562
Db 529 AGGAAGCCAGACTCAGGAAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 588
QY 563 TTTTCAACATCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 622
Db 589 AGACCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 648
QY 623 ATGCTATGATATGACCTTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 682
Db 649 ACCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 708
QY 683 TTTTCAAAATACCTTACTGAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 742
Db 709 TCTACAAGAGGCAAGAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 768
QY 743 ACTATTGGAAGAAATGAGCCCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 802
Db 769 ACAGTGGCTGCTGAGAGGGGACCCCTGCCAGCAAGCTGATCTTGGCATGCTGCTGCTGCTGCT 828
QY 803 GACACACCTTCTATCTGAGAAACCCCTCTGATGATGATGATGATGATGATGATGATGATGATG 862
Db 829 GAGCTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 888
QY 863 ATGGCCCTGCTGGCGCTTATACAGACAGCTGGGTCTGGGCTTCTGCTGCTGCTGCTGCTGCTGCT 922
Db 889 CTGGCACTCCAGGCCCTTCAACAGAGGAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 948
QY 923 CTTTCTGAGAGTGGAGCCACTGAGGTCTGGGATGCTGCCATCCCAAGAGTGGCCCTATGCTGCT 982
Db 949 CCT-----GGAAGGGGGCCACCAACAGAGATCCAGGATCAGAGGTGGCCCTACATCT 1002
QY 983 ATAAAGCCACAGTGGCTTGGCTATGACAAATATCAAGAGCTTCAAGTGTAAAGCTCAGT 1042

Db 1003 TCCGGGACACACAGTGGGTGGCTTTGATGATGTGGAGAGCTTCAAAACCAAGGTGAGCT 1062
QY 1043 GCCTTAAGCAGAACAAATTTGGAGTGCCATGATCTGGCCATTGACCTTGATGACTTCA 1102
Db 1063 ATCTGAAGCAGAAAGGAGCTGGGGGGGCGCATGCTCTGGGCACCTGGGACTTAGATGACTTTG 1122
QY 1103 CTGGCTCTTCTGTGATGACGGGAAATTTCCCTCTGA 1138
Db 1123 CCGGCTCTCTGCAACACGAGGCGCATACCCCTCA 1158

RESULT 12
US-09-267-574-3
; Sequence 3, Application US/09267574
; Patent No. 639571
; GENERAL INFORMATION:
; APPLICANT: Gray, Patrick W.
; APPLICANT: Tjoelker, Larry W.
; TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
; FILE REFERENCE: 27866/35407
; CURRENT APPLICATION NUMBER: US/09/267,574
; CURRENT FILING DATE: 1999-03-12
; EARLIER APPLICATION NUMBER: 09/039,198
; EARLIER FILING DATE: 1998-03-12
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1656
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (27)..(1424)
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (90)..(1424)
US-09-267-574-3

Query Match 27.9%; Score 425.2; DB 4; Length 1656;
Best Local Similarity 62.0%; Pred. No. 2.1e-122;
Matches 692; Conservative 0; Mismatches 418; Indels 6; Gaps 1;
QY 23 CAGGTCTGCTCTTCTGCTGAATGCTCAGCTGGGCTCTGCTACAATCTGATGCTATT 82
Db 49 CAGGTTTCATGCTCTGCTGATGATCCATGGGCTCTGCTGCAAACTGGTCTGCTACT 108
QY 83 TCACCAACTGGGCCCCAGTATCGCCAGCTCTGGGAGCTTCAAGCCTGATGATTAACC 142
Db 109 TCACCAACTGGGCCCCAGTATCGCCAGCTCTGGGAGCTCTGCTTCTTCCGCAAGGACTTGGACC 168
QY 143 CTTGCTGTGCTACTACCTGATCTATGCTTCTGCTGGATGAGAACAAATGAGATCACCA 202
Db 169 CAGCCTTTCGACCCACCTCTCTAGCCTTCTGCTGCTGATGACCAACCCAGCTGAGCA 228
QY 203 CCATAGAATGGAATGATGTTACTCTCTATAAGCTTTCAATGACTTGAAAGAACAGGAACA 262
Db 229 CCACTGATGGAATGACGAGACTCTCTACCAGGAGTCAATGCTTGAAGAAAGATGATC 288
QY 263 CCAACTGAAACCCCTCTGGCAATTTGGAGCTGGAACTTGAAGTCTTCTTCACTA 322
Db 289 CCAAGCTGAGAACCCCTGTTAGCCATCGAGGCTGGAATTTGACACTCAGAACTTCACAG 348
QY 323 CCATGTTTCCACTCTCAGAACCCGACAGCTTCTTACTCTCAGTCACTCAAAATTTCTGC 382
Db 349 ATATGTAGCACCGGCAACACCGTACAGCTTTGTCAACTCGGCATCAGGTTTCTGC 408
QY 383 GTCAGTATGGTTTGTGACCTGGACCTGGACTGGGAATACCCAGGCTCAGCTGGGAGCC 442
Db 409 GCAATACAGCTTTGACGGCTTGACCTTGAGCTGGGAGTACCCAGGAGCCAGGGAGCC 468
QY 443 CTCTCAGGACAGATCTCTTCACTGCTCTGCTGAGGAAATGCGTGAAGCTTTTTCAGC 502

Db 469 CTCCGTAGACAAAGGAGCGCTTCAACACCTGTACAGGACTTTGGCCAATSCCTTCCAGC 528
QY 503 AGGAGGCTATTGAGAGCAACAGAGCCAGAGCTGATGTTACTGCTGTGTAGCTGGTGGGA 562
Db 529 AGGAGCCAGAGCTCAGGAAGAGCGCTTCTTCTGAGTGAGCGGTTCCAGCTGGGC 588
QY 563 TTTTCAACATPCCAGGCTGCTATGAGATCCCTGAACCTTTCTAAGTACCTGGATTTTCATCC 622
Db 589 AGACCTATGTGGATGCTGGATAGAGGTGGACAAAATCGCCAGAACCTGGATTTTGTCA 648
QY 623 ATGTCATGACATATGACCTCCATGGCTCTCTGGAGGCTACATGCGGGAAGTATGCTC 682
Db 649 ACCTTATGGCTTACGACTTCCATGGCTCTTGGAGAGGTCAAGGGGACATAACAGCCCCC 708
QY 683 TTTTACAAATACCTTACTGAGACTGGTAGCAATGCTACCTCAATGTGGATTTATGTCATGA 742
Db 709 TCTACAGAGGCAAGAGAGAGTGGTGACAGCCAGCCCTCAACGTGGATGCTGCTGC 768
QY 743 ACTATTGGAAGAACAAATGGAGCCCCAGCTGAGAGCTCATTTGTTGGATTCCCAGAGATG 802
Db 769 AACAGTGGCTGACAGAGGGGAGCCCTGCCAGCAAGCTGATCTTGGCATGCCCTACCTACG 828
QY 803 GACACACCTTCACTCTGAGAAACCCCTCTCATTAATGGAATTTGGTGGCCCTTACCTCTG 862
Db 829 GAGGCTCTCTACACTGGCTCTCATACAGACACAGAGTGGGGGCCCCAGCCAGGGT 888
QY 863 ATGGCCCTGCTGGCGCTATACAGACAGAGCTGGGTTCTGGGCTACTATGAGATTTGCA 922
Db 889 CTGGCACTCAGGCCCTTCCACAGGAAGAGGATGCTGGCTTACTATGAAGTCTGCT 948
QY 923 CTTTCTGAGAAGTGGAGCCACTGAGTCTGGATGCTCTCCCAAGAGTCCCTTATGCCCT 982
Db 949 CCT-----GGAAGGGGCCCAACAGAGAAATCCAGGATCAGAGGTGCCCTTACATCT 1002
QY 983 ATAAGGCCAAGAGTGGCTTGGCTATGACAAATATCAAGAGCTTCAGTGTTAAGGCTCAGT 1042
Db 1003 TCCGGGACAAACAGTGGGTGGGCTTTGATGATGTGGAGAGCTTCAAAACCAAGGTGAGCT 1062
QY 1043 GGCTTAAGCAGAAATTTTGGAGTGCCATGATCTGGGCAATGACCTTGTATGACTTCA 1102
Db 1063 ATCTGAAGCAGAGGAGCTGGGGGGCCATGCTTGGGCACCTGGGACTTGAAGTACTTTG 1122
QY 1103 CTGCGCTTCTCTGATGATCAGGAAATTTTCTCTCTGA 1138
Db 1123 CCGGCTCTCTGCAACACGAGGCGCATACCCCTCA 1158

RESULT 13
US-08-694-915-1
; Sequence 1, Application US/08694915
; Patent No. 5811535
; GENERAL INFORMATION:
; APPLICANT: Adamou, Julie
; APPLICANT: Kirkpatrick, Robert
; APPLICANT: Rosenberg, Martin
; TITLE OF INVENTION: HUMAN CARTILAGE GP39-LIKE GENE
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Smithkline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-2799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA: US/08/694,915
; APPLICATION NUMBER: US/08/694,915
; FILING DATE:
; CLASSIFICATION:

INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1526 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE:
US-08-694-915-3

Query Match 21.7%; Score 330.8; DB 1; Length 1526;
Best Local Similarity 58.9%; Pred. No. 5.5e-93;
Matches 657; Conservative 0; Mismatches 432; Indels 27; Gaps 4;

QY 55 GGGCTGCTCAATCTGATATGCTATTTTACCAACTGGGCCAGTATCGGCCAGGTCTG 114
DB 199 GGATCTGGCTTCAAACTGGTTTGTACTTTACCAACTGGTCCAGGACCGGCAAGCA 258
QY 115 GGGAGCTTCAAGCTGATGACATTAACCCCTGCTGTCTACTCACTGATCTATGCTTT 174
DB 259 GGAATTCACCCCTGAGAATATTGACCCCTTCTCTATGCTCTCATCTCTATTCATTC 318
QY 175 GCTGGGATCGAGAACAATGAGATCACCACCATAGAAATGGAATGATGTTACTCTCTATAA 234
DB 319 GCCAGCATCGAACAACAAGTATATCATCAAGGACAGAGTGAAGTATGCTCTACAG 378
QY 235 GCTTTCATGACTTGAAGAACAGGACAGCAAACTGAAACCCCTCTGGCAATGGAGGC 294
DB 379 ACCATCAACAGCTCAAAACCAAGAAATCCCAAACTGAAATTTCTCTGCTCCATTTGGAGG 438
QY 295 TGGAACTTTGAACTGCTCTTCACTACCATGTTTCCACTTCTCAAAACCGCCAGACC 354
DB 439 TACCTGTTTGTTCAAAGGGTTCACCCCTATGTTGGATTTCTTCATCATCAGCTTGAA 498
QY 355 TTCATTACCTCAGTCATCAAAATTTCTGCTGATGTTGTTGATGGACTGGACCTGGAC 414
DB 499 TTCATTAACTCCAAATCCTCTTCTGAGGAACCAATACTTTGATGGACTGATGTAAGC 558
QY 415 TGGNAATACCAAGCTCAGTGGGAGCCCTCTCAGGACAGATCTCTTCACTGCTCTG 474
DB 559 TGGATCTACCCAGA-----TCAGAAAGAAACACTCATTTTCACTGCTGCTG 603
QY 475 GTGAAGAAATGCGTGAAGCTTTTGAAGCAGAGGCTATTGAGAGCAACAGCCAGACTG 534
DB 604 ATTCATGATTTAGCAGAACCCCTTTCAGAAAGACTTCAAAATCCCAAGAAAGGCTT 663
QY 535 ATGTTTACTGCTGTAGCTGAGTGGGATTTTCCAACTCCAGGCTGGCTATGATCCCT 594
DB 664 CTCCTGACTCGGGGCTATCTGCGAGGAGCAATGATTGATAACAGCTATCAAGTTGAG 723
QY 595 GAACCTTCTAGTACCTGGATTTATCATGTCATGATGATGATGATGATGATGATGATG 654
DB 724 AAACCTGGAAAGATCTGGATTTATCATCAACCTCTGCTCTTGTGCTTCCATGGTCTGG 783
QY 655 GA-----GGGCTACATCGGGAGAAATAGTCTCTTTTACAATAACCTTACTGAGACTGT 708
DB 784 GAAAGGCCCTTATCAGTGGCCACACAGCCCTCTGAGCAAGGGGTGGCAGACAGAGGG 843
QY 709 AGCAATGCTACCTCAATGTGGATTTATGATGAATTTGGAAGAAACAATGGAGCCCA 768
DB 844 CCAAGCTCTACTACAATGTGAATATGCTGTGGGTACTGATATACATATGGAATGCA 903
QY 769 GCTGAGAGCTCATTGTTGGATTCGAGATGATGACACACCTTCACTCTGAGAAACCC 828
DB 904 TCAGAAAGGTGGTCAATGGGATGCCCCACATATGGGCACTCTCTCACACTG---GCCCT 960
QY 829 TCTGATATGAAATTTGGTGGCCCTTACCTCTGATGCTGCTGCTGCTGCTGCTGCTGCT 888
DB 961 GCAGAAACCCCGTGGGGGCCCTTGCCTCTGGCCCTGGAGCTGTGCTGCTGCTGCTGCTGCT 1020

QY 889 CAGGCTGGGTTCTGGGCCCTACTACTATGAGATTTGCACCTTTCTGAGAAAGTGGAGCCACTGAG 948
DB 1021 TCTTCAGGCTTCTGGCCCTATTATGAGATTCGCCAGTTCCTGAAA---GGAGCCCAAGATC 1077
QY 949 GTCTGGGATGCTCCCAAGAGTGGCCCTATGCTATAGGCAACGAGTGGCTTGGCTAT 1008
DB 1078 ACGCGGCTCCAGGATCAGCAGGTTCCCTACGAGTCAAGGGGAACAGCTGGGTGGCTAT 1137
QY 1009 GACAATATCAAGAGCTTCAGTGTAAAGCTCAGTGGCTTAAGCAAGCAAAATTTTGGAGGT 1068
DB 1138 GATGATGTAAGAGTATGGAGCAAGGTTCACTTCTTAAGAAATTTAAACCTGGAGGA 1197
QY 1069 GCAATGATCTGGGCCATTCACCTTGATGACTTCACTGCTCTTTCTGATGATCAGGAAAA 1128
DB 1198 GCAATGATCTGCTTATTTGACATGATGACTTCACTTGGCAAACTCTGCAACCAAGGCCCT 1257
QY 1129 TTTCTCTGACTTCTACTTTTGAACAAAGCCCTTGGC 1164
DB 1258 TACCTCTTGTCCAAAGCAGTCAAGAGAGCCCTTGGC 1293

RESULT 15

PCT-US94-07754-4
; Sequence 4, Application PC/TUS9407754
; GENERAL INFORMATION:
; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
; TITLE OF INVENTION: ASSAY FOR YKL-40 AS A MARKER FOR
; TITLE OF INVENTION: DEGRADATION OF MAMMALIAN CONNECTIVE TISSUE MATRICES
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SPENSLEY HORN JUBAS & LUBITZ
; STREET: 1800 CENTURY PARK EAST, FIFTH FLOOR
; CITY: LOS ANGELES
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/07754
; FILING DATE: 08-JUL-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: HOWELLS, STACY L.
; REGISTRATION NUMBER: 34,842
; REFERENCE/DOCKET NUMBER: FD 3665
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/455-5100
; TELEFAX: 619/455-5110
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1681 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: YKL-40
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 135..1681
PCT-US94-07754-4

Query Match 20.6%; Score 314.8; DB 5; Length 1681;
Best Local Similarity 58.2%; Pred. No. 5.7e-88;
Matches 641; Conservative 0; Mismatches 437; Indels 24; Gaps 4;
QY 22 ACAGTCTGGCTCTTCTGCTGAATGCTCAGCTGGGGTGTGCTTACATCTGATGCTAT 81
DB 93 ACAGCTTTTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 152

Search completed: July 3, 2003, 08:43:20
Job time : 75.5873 secs

```
QY 82 TTCACCAACTGGGCCAGTATCGCCAGAGTCTCTGGGAGCTTCAAGCCTGATGACATTAAAC 141
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 153 TACACAGCTGGTCCAGTACCGGAGGAGGAGGAGCTGCTCCAGATGCCCTTGAC 212
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 142 CCTGCTGTGTACTACCTGATCTATGCCCTTGTGGGATGCGAGAACAAATGAGATCAC 201
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 213 CGCTTCTGTGTACCCACATCATCTACAGCTTTGCCAATATAAGCAACGATCACATCGAC 272
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 202 ACCATAGAATGGAATGATGTTACTCTCTATAAAGCTTTCATGACTTGRAAAACAGGAAC 261
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 273 ACCTGGAGTGGAAATGATGAGCGCTCTACGGCATGCTCAACACACTCAACACACGAAC 332
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 262 AGCAAACTGAAACCTCTCTGGCAATTTGGAGCTTGAACCTTTTGAACCTGCTCTCTCACT 321
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 333 CCCAACTGAAGACTCTCTGTCTGCGGAGGATGGAACCTTTGGGTCTCAAGATTTTCC 392
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 322 ACCATGGTTTCCACTTCTCAGAACCGCCAGACCTTCATTTACCTCAGTCATCAATTTCTG 381
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 393 AAGATAGCCTTCCAAACCCAGAGTTCGCGGACTTTTCATCAAGTCAGTACCGCCATTTCTG 452
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 382 CGTCAGTATGGGTTTATGAGCTGGACCTGGAGTACCCAGGCTCACGTGGGAGC 441
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 453 CGCACCCATGGCTTTGATGGCGTGACCTGCTGCTGCTTACCTGGACGGAGA----- 506
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 442 CTCTCTCAGACAAGCATCTCTTCACTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 501
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 507 -----GACAAACACCATTTTACCACCTTAATCAAGGAATGAAGGCCGAATTTATA 557
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 502 CAGGAGGCTATTGAGAGCAACAGGCCAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 561
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 558 AAGGAAGCCAGCCAGCGGAAAAAGC---AGCTCTCTCAGCGCAGCAGCTGTCTGCGGGG 614
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 562 ATTTCCAACATCCAGGCTGGCTATGAGTCCCTGAACTTTCTTAAGTACCTGGATTTTCATC 621
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 615 AAGTCAACATTCACAGCAGCTATGATGCTCAAGATATCCCAACACCTGGATTTCAAT 674
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 622 CATGTCATGACATATACCTTCCATGCTCTGCGGAGGCTTACACTGGGAGAAATAGTCTCT 681
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 675 AGCATCATGCTACCATTTTCTGCGCTGGGCTGGGACACAGGCCATCACAGTCCC 734
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 682 CTTTACAATACCTTACTGAGCTGTAGCAATGCTTACCTCAATGTGGATATATGTCATG 741
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 735 CTGTTCCGAGGTCAGGAGTCAAGTCTGACAGATTGAGCAACACTGACTATGCTGTG 794
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 742 AACTATTGGAAGAACAAATGAGGCCAGCTGAGAGCTCATTTGTTGATTCCACAGAGTAT 801
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 795 GGGTACATGTTGAGGCTGGGGCTCTGCCAGTAAGCTGGTATGGGCAATCCCACTTC 854
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 802 GGACACACCTTCACTCTGAGAAACCCCTCTGATTAATGGAATGGTGGCCCTACCTCTGTT 861
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 855 GGGAGGAGCTTCACTCTG---GCTTCTTCTGAGACTGGTGTCCAGCGCAATCTCAGGA 911
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 862 GATGGCCCTGCTGGCCCTTATACCAGACAGGCTGGGTTCTGGGCCCTACTATGAGATTTGC 921
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 912 CGGGAAATTCACGGCCGGTTACCAAGGAGGAGGAGCCCTTGCCTACTATGAGATCTGT 971
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 922 ACCTTCTCAGAAAGTGAGCCTAGGCTGGGATGCTCCCAAGAGTGCCCTATGCC 981
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 972 GACTTCT---CGCGGAGCCACAGTCCATAGAACCTCGGCAGCAGGTCCCTATGCC 1028
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 982 TATAAGGCCAAGAGTGGCTTGGCTATGACAATATCAAGAGCTTCAAGTGTAAAGGCTCAG 1041
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1029 ACCAAGGGCAACAGTGGGTAGATACGACGACAGGAAAGCGTCAAAAGCAAGGTGCAG 1088
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1042 TGGCTTAAGCAGACAAATTTGGAGTGCCATGATCTGGGCCATTTGACCTTGTGACTTC 1101
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1089 TACCTGAAGGATAGGAGCTGGCAGGCCCATGGTATGGGCCCTGGACCTGGATGACTTC 1148
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1102 ACTGGCTCTTCTGTGATCAGG 1123
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1149 CAGGGCTCTTCTGCGGCCAGG 1170
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: July 3, 2003, 05:03:26 ; Search time 267.722 Seconds
(without alignments)
8891.723 Million cell updates/sec

Title: US-10-004-219b-3
Perfect score: 1525
Sequence: 1 atggcgaagctactcttctgt.....taaaattgtagccaaaca 1525

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1085931 seqs, 780495707 residues
Total number of hits satisfying chosen parameters: 2171862

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Published_Applications_NA.*
- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
 - 2: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
 - 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
 - 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
 - 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
 - 6: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
 - 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
 - 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
 - 9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
 - 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
 - 11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
 - 12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
 - 13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
 - 14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1525	100.0	1525	9	US-10-004-219b-3
2	1041	68.3	1625	9	US-10-004-219b-3
3	430	28.2	1038	9	US-10-202-436A-14
4	327.6	21.5	1391	10	US-09-823-830A-402
5	327.6	21.5	449	10	US-09-960-352-678
6	319.6	21.0	1925	9	US-10-097-340-44
7	311.6	20.4	1681	10	US-09-215-077A-4
8	311.6	20.4	1681	10	US-09-262-213A-4
9	259.8	17.0	1474	10	US-09-765-231A-8
10	152.2	10.0	415	10	US-09-960-352-2589
11	149.6	9.8	414	10	US-09-960-352-13050
12	148	9.7	410	10	US-09-960-352-7364
13	148	9.7	424	10	US-09-960-352-3057
14	147.4	9.7	418	10	US-09-960-352-3072
15	136.2	8.9	399	10	US-09-960-352-265
16	134.8	8.8	411	10	US-09-960-352-14919
17	127.8	8.4	384	10	US-09-960-352-5721
18	124.6	8.2	398	10	US-09-960-352-14460
19	121.4	8.0	1608	9	US-10-218-743-20

c	20	121.4	8.0	1608	9	US-10-218-743-22	Sequence 22, Appl
c	21	121.4	8.0	1665	9	US-10-218-743-17	Sequence 17, Appl
c	22	121.4	8.0	1865	9	US-10-218-743-19	Sequence 19, Appl
c	23	121.4	8.0	1752	9	US-10-218-743-14	Sequence 14, Appl
c	24	121.4	8.0	1752	9	US-10-218-743-16	Sequence 16, Appl
c	25	112.8	7.4	1470	9	US-10-218-743-40	Sequence 40, Appl
c	26	112.8	7.4	1470	9	US-10-218-743-42	Sequence 42, Appl
c	27	112.8	7.4	1527	9	US-10-218-743-37	Sequence 37, Appl
c	28	112.8	7.4	1527	9	US-10-218-743-39	Sequence 39, Appl
c	29	112.8	7.4	1621	9	US-10-218-743-34	Sequence 34, Appl
c	30	112.8	7.4	1621	9	US-10-218-743-36	Sequence 36, Appl
c	31	97.8	6.4	373	9	US-09-918-995-8214	Sequence 8214, Ap
c	32	91.2	6.0	437	9	US-09-918-995-35463	Sequence 35463, A
c	33	83	5.4	416	9	US-09-918-995-6770	Sequence 6770, Ap
c	34	80.2	5.3	181	9	US-09-796-692-9551	Sequence 9551, Ap
c	35	80.2	5.3	181	9	US-10-040-862-9551	Sequence 9551, Ap
c	36	79.8	5.2	438	10	US-09-960-352-1882	Sequence 1882, Ap
c	37	79.2	5.2	423	10	US-09-960-352-12100	Sequence 12100, A
c	38	78.8	5.2	430	10	US-09-960-352-1189	Sequence 1189, Ap
c	39	78.2	5.1	422	10	US-09-960-352-5957	Sequence 5957, Ap
c	40	76.6	5.0	422	10	US-09-960-352-10642	Sequence 10642, A
c	41	76.2	5.0	419	10	US-09-960-352-4053	Sequence 4053, Ap
c	42	76.2	5.0	419	10	US-09-960-352-10172	Sequence 10172, A
c	43	75.2	4.9	426	10	US-09-960-352-4979	Sequence 4979, Ap
c	44	74.8	4.9	410	10	US-09-960-352-8317	Sequence 8317, Ap
c	45	74.8	4.9	440	10	US-09-960-352-4058	Sequence 4058, Ap

ALIGNMENTS

RESULT 1
US-10-004-219b-3
; Sequence 3, Application US/10004219B
; Publication No. US20030087414A1
; GENERAL INFORMATION:
; APPLICANT: Macrozyme
; APPLICANT: Aerts, Johannes M.F.G.
; APPLICANT: Root, Rolf G.
; TITLE OF INVENTION: A mammalian mucinase, its recombinant production, and
; TITLE OF INVENTION: Its use in therapy or prophylaxis against diseases in
; TITLE OF INVENTION: Which mucus is involved or infection diseases
; FILE REFERENCE: 2183-5136US
; CURRENT APPLICATION NUMBER: US/10/004,219B
; CURRENT FILING DATE: 2001-11-02
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1525
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Combined DNA/RNA Molecule: mouse
; OTHER INFORMATION: AMCase cDNA sequence and deduced amino acid
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: mouse AMCase
; OTHER INFORMATION: cDNA sequence and deduced amino acid sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1419)
US-10-004-219b-3

Query Match 100.0%; Score 1525; DB 9; Length 1525;
Best Local Similarity 100.0%; Pred: No. 0;
Matches 1525; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCGAAGCTACTTCTCGTCACAGGCTGGGCTTCTTCTGCTGAATGCTCAGCTGGGGTCT 60
Db 1 ATGGCGAAGCTACTTCTCGTCACAGGCTGGGCTTCTTCTGCTGAATGCTCAGCTGGGGTCT 60
QY 61 GCGTCAAACTGATGCTATTTCACCAACTGGGCCAGTATCGGCCAGGCTCTGGGGAGC 120
|||||

Db 61 GCCTACATCTGATATGCTATTTCCACCACTGGGCCAGTATCGGCCAGGTCTGGGAGC 120
QY 121 TTCAAGCTGATGACATTAACCCCTGCCTGTGTACTACCTGATCTATGCTTTGCTGGG 180
Db 121 TTCAAGCTGATGACATTAACCCCTGCCTGTGTACTACCTGATCTATGCTTTGCTGGG 180
QY 181 ATCAGAACATGAGATCACCACCATAGAAATGGAATGATTTACTCTCTATAAAGCTTTC 240
Db 181 ATCAGAACATGAGATCACCACCATAGAAATGGAATGATTTACTCTCTATAAAGCTTTC 240
QY 241 AATGACTTGAAGAACAGGACAGCAAACTGAAACCCCTCTGCAATTTGGAGGCTGGAAC 300
Db 241 AATGACTTGAAGAACAGGACAGCAAACTGAAACCCCTCTGCAATTTGGAGGCTGGAAC 300
QY 301 TTGGAACCTGCTCTTTCACTACCATGGTTTCCACTTCTCAGAACCCGACAGCTTCATT 360
Db 301 TTGGAACCTGCTCTTTCACTACCATGGTTTCCACTTCTCAGAACCCGACAGCTTCATT 360
QY 361 ACCTCAGTCAATCAAAATTTCTGCGTCAGTATGGGTTTGTAGGACTGGACCTGGGAA 420
Db 361 ACCTCAGTCAATCAAAATTTCTGCGTCAGTATGGGTTTGTAGGACTGGACCTGGGAA 420
QY 421 TACCCAGGCTCAGTGGGAGCCCTCTCAGGACAAGCATCTTCACTGTCTGTGGTGAAG 480
Db 421 TACCCAGGCTCAGTGGGAGCCCTCTCAGGACAAGCATCTTCACTGTCTGTGGTGAAG 480
QY 481 GAAATCGTGAACCTTTGAGCAGGAGGCTATTGAGAGCAACAGGCCAGACTGATGGTT 540
Db 481 GAAATCGTGAACCTTTGAGCAGGAGGCTATTGAGAGCAACAGGCCAGACTGATGGTT 540
QY 541 ACTGCTGTGACTGCTGGGATTTCCACATCCAGGCTGGCTATGAGATCCCTGAACCT 600
Db 541 ACTGCTGTGACTGCTGGGATTTCCACATCCAGGCTGGCTATGAGATCCCTGAACCT 600
QY 601 TCTAAGTACTGATTCATCCATGTCATGACATATGACCTCCAGTCCCTGGGAGGC 660
Db 601 TCTAAGTACTGATTCATCCATGTCATGACATATGACCTCCAGTCCCTGGGAGGC 660
QY 661 TACACTGGGAGAAATAGTCTCTTTACAAATACCCTACTGAGACTGGTATGATGCTAC 720
Db 661 TACACTGGGAGAAATAGTCTCTTTACAAATACCCTACTGAGACTGGTATGATGCTAC 720
QY 721 CTCATGTGGATTCATGATGAACTATTGGAAGAACAAATGGAGCCCGAGCTGAAAGCTC 780
Db 721 CTCATGTGGATTCATGATGAACTATTGGAAGAACAAATGGAGCCCGAGCTGAAAGCTC 780
QY 781 ATTGTGGATTCACAGATGACACACCTTCATCTGAGAACCCCTCTGATATGGA 840
Db 781 ATTGTGGATTCACAGATGACACACCTTCATCTGAGAACCCCTCTGATATGGA 840
QY 841 ATGTTGCCCTACCTCTGGTATGGCCCTGCTGGCCCTATACAGACAGGCTGGGTTTC 900
Db 841 ATGTTGCCCTACCTCTGGTATGGCCCTGCTGGCCCTATACAGACAGGCTGGGTTTC 900
QY 901 TGGGCTACTATGAGATTTGACCTTTCTGAGAAAGTGGAGCCACTGAGGTCTGGATGCC 960
Db 901 TGGGCTACTATGAGATTTGACCTTTCTGAGAAAGTGGAGCCACTGAGGTCTGGATGCC 960
QY 961 TCCCAAGAGTGCCTTCATGCTATAAGGCCAAGCTGGCTTGGCTATGACAAATATCAAG 1020
Db 961 TCCCAAGAGTGCCTTCATGCTATAAGGCCAAGCTGGCTTGGCTATGACAAATATCAAG 1020
QY 1021 AGCTTCAGTGTAAAGCTCAGTGGCTTAAAGCAGAACAAATTTTGGAGGTGCCATGCTGG 1080
Db 1021 AGCTTCAGTGTAAAGCTCAGTGGCTTAAAGCAGAACAAATTTTGGAGGTGCCATGCTGG 1080
QY 1081 GCCATTGACCTTGATGACTTCACTGGCTTCTGTGATCAGGGAATTTTCTCTGACT 1140
Db 1081 GCCATTGACCTTGATGACTTCACTGGCTTCTGTGATCAGGGAATTTTCTCTGACT 1140
QY 1141 TCTACTTTTGAACAAAGCCCTTGGCATATCCACTGAAAGGTTGACACGCTCTGACGTGCCT 1200
Db 1141 TCTACTTTTGAACAAAGCCCTTGGCATATCCACTGAAAGGTTGACACGCTCTGACGTGCCT 1200

QY 1201 TCCGAGCCAGTGAATCTCTCTCCAGGAAGTGGGAGTGGGGTGAAGCTCCCGAGGAAGC 1260
Db 1201 TCCGAGCCAGTGAATCTCTCTCCAGGAAGTGGGAGTGGGGTGAAGCTCCCGAGGAAGC 1260
QY 1261 TCTGGAGGAGTGAATCTCTCTCCGAGCAAGAGAGATGGCTCTACCTCTGGCAGATGAC 1320
Db 1261 TCTGGAGGAGTGAATCTCTCTCCGAGCAAGAGAGATGGCTCTACCTCTGGCAGATGAC 1320
QY 1321 AGAAATGCTTTTGGCAGTGCATCAATGGAATCACATACCAGCAGCATTTCAAGCAGGG 1380
Db 1321 AGAAATGCTTTTGGCAGTGCATCAATGGAATCACATACCAGCAGCATTTCAAGCAGGG 1380
QY 1381 CTGTTTTGATACCAAGCTGTAAATGCTGCAACTGGCCATGAACCTTAATGCCATTTCTCC 1440
Db 1381 CTGTTTTGATACCAAGCTGTAAATGCTGCAACTGGCCATGAACCTTAATGCCATTTCTCC 1440
QY 1441 AGAAATTTCTGCACTCTCTTTACTCTCTACCAAAAGTAACTATCTCCCTTTAACTTA 1500
Db 1441 AGAAATTTCTGCACTCTCTTTACTCTCTACCAAAAGTAACTATCTCCCTTTAACTTA 1500
QY 1501 TGCAATAAAATTTGGTAGCCAAAACA 1525
Db 1501 TGCAATAAAATTTGGTAGCCAAAACA 1525

RESULT 2

US-10-004-219b-2
; Sequence 2, Application US/10004219B
; Publication No. US2003008741A1
; GENERAL INFORMATION:
; APPLICANT: Macrozyme
; APPLICANT: Aerts, Johannes M.F.G.
; APPLICANT: Boot, Rolf G.
; TITLE OF INVENTION: A mammalian mucinase, its recombinant production, and its use in therapy or prophylaxis against diseases in which mucus is involved or infection diseases
; FILE REFERENCE: 2183-5136US
; CURRENT APPLICATION NUMBER: US/10/004,219B
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1625
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Combined DNA/RNA Molecule: human
; OTHER INFORMATION: AMCase cDNA sequence and deduced amino acid
; OTHER INFORMATION: sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: human AMCase
; OTHER INFORMATION: cDNA sequence and deduced amino acid sequence
; NAME/KEY: CDS
; LOCATION: (104)..(1531)
US-10-004-219b-2

Query Match 68.3%; Score 1041; DB 9; Length 1625;
Best Local Similarity 83.4%; Pred. No. 0;
Matches 1198; Conservative 0; Mismatches 230; Indels 9; Gaps 1;
QY 1 ATGCCCAAGCTACTTCTCTGTCACAGGTCTGGCTTCTTCTGCTGAATGCTCAGCTGGGCTCT 60
Db 104 ATGCAAAAGCTTATTCTCTCTCAGGTCTTGCTTTATCTGAATTTGCGCTCGGCTCT 163
QY 61 GCCTACATCTGATGCTATTTCACCAACTGGGCCAGCATGTCGGCCAGCTCTGGGAGC 120
Db 164 GCCTACAGCTGACATGCTACTTCACCAACTGGGCCAGTACCGGCCAGGCTCTGGGCGC 223
QY 121 TTCAAGCTTGATGACATTAACCCCTGCCTGTGTACTGACCTGATCTATGCTTTGCTGG 180
Db 224 TTCTATGCCCTGACAAACATGACCCCTGCTCTGTACCCACCTGATCTACGCTTTGCTGG 283

Qy	181	ATGCAGAACAAATGAGATCACCCATAGAAATGAATGATTTACTCTCTATAAAGCTTTC	240
Db	284	AGGCAGAACAAACGAGATCACCCATCGAATGGAAGGATGTGACTCTCTACCAAGCTTTC	343
Qy	241	AATGACTTGAAAACAGGAACACCAACTGAAAACCCCTCCTCGCAATTGGAGGCTGGAAC	300
Db	344	AATGGCCTGAAAATATAGAACACGACGCTGAAAACCTCTCTGCCCATTGGAGGCTGGAAC	403
Qy	301	TTTGGAACTGCTCTTTCACTACCATGTTTCCACTTCTCAGAACGCCACAGACTTCATT	360
Db	404	TTGGGACTGCCCTTTCACTGCCATGGTTTCTACTCTCGAAGACGCCAGACTTTCATC	463
Qy	361	ACCTCAGTCATCAAAATTTCTGGCTCAGTATGGGTTTGTATGGACTGACCTGGAGCTGGAA	420
Db	464	ACCTCAGTCATCAAAATTTCTGGCCCAAGTATGAGTTTGAACGGGTGGACTTTGACTGGGAG	523
Qy	421	TACCAGGCTCAGTGGGAGCCCTCTCAGGACAAAGCATCTCTTCACTGTCTCTGGTGAAG	480
Db	524	TACCTGGCTCTCGTGGAGCCCTCTCAGGACAAAGCATCTCTTCACTGTCTCTGGTGCAG	583
Qy	481	GAATGGCTGAAAGCTTTTGAGCAGGAGGCTATTGAGAGCAACAGGCCACAGACTGATGGTT	540
Db	584	GAATGGCTGAAGCTTTTGAGCAGGAGGCCAAGCAGATCAACAAGCCACGGCTGATGGTC	643
Qy	541	ACTGCTGCTGTAGCTGGTGGGATTTTCCAAACATCCAGGCTGGCTATGAGATCCCTGAACTT	600
Db	644	ACTGCTGCAGTAGCTGCTGGCACTCTCCAATATCCAGTCTGGCTATGAGATGCCCAACTG	703
Qy	601	TCTAAGTACCTGGATTTTCATCTCATGATGACATATGACCTCCATGCTCCTGGGAGGGC	660
Db	704	TCACAGTACCTGGACTACATCCATGTCTATGACCTACGACCTCCATGGCTCCTGGGAGGGC	763
Qy	661	TACACTGGGAGAAATAGTCTCTTTCAAAATACCTTACTCAGACTGGTATGCAATAGCCTAC	720
Db	764	TACACTGGAGAACACGCCCCCTCTCAAAATACCCGACTCACACCGGACGACCAACGCCTAC	823
Qy	721	CTCAATGTGATATATGTCATGAATAATTGGAAGAACAAATGAGCCCCAGCTGACAGCTC	780
Db	824	CTCAATGTGATATATGTCATGAATAATTGGAAGAACAAATGAGCCCCAGCTGACAGCTC	883
Qy	781	ATTGTTGGATTCCCAGATATGGACACACCTTCATCTGAGAAACCCCTCTGATAATGGA	840
Db	884	ATCGTTGGATTCCCCTACTTATGGACACAACTTCATCTGAGCAACCCCTCCAACACTGGA	943
Qy	841	ATTTGGTCCCTTACTCTGTGTATGGCCCTCTGTGGCCGCTATACCAACAGGCTGGGTTTC	900
Db	944	ATTTGGTCCCTTACTCTGTGTGTGGTCTGTCTGTGGCCCTATGCAAGGAGTCTGGGATC	1003
Qy	901	TGGGCTTACTATGAGATTTGCACCTTCTCAGAAAGTGGACCACTGAGGTCTGGGATGCC	960
Db	1004	TGGGCTTACTACGAGATCTGTACCTTCTGAAAATGAGCCACTCAGGATGGGATGCC	1063
Qy	961	TCCCAAGAAGTGCCCTATGCTTATAGGCCCAACAGTGGCTTGGCTATGACAATATCAAG	1020
Db	1064	CCTCAGGAAGTGCCCTTATGCTTATCAGGGCAATGTGTGGTGGCTTGGCTATGACAATCAAG	1123
Qy	1021	AGCTTCAGTGTTAAGGCTCAGTGGCTTAAAGCAGAACAAATTTTGGAGGTGCCATGATCTGG	1080
Db	1124	AGCTTCGATATTAAGGCTCAATGGCTTAAAGCAGAACAAATTTTGGAGGCGGATCTGGT	1183
Qy	1081	GCCATTGACCTTGATGACTTCACTGGCTCTTTCTGTGTCATCAGGGAATAATTCCTCTGACT	1140
Db	1184	GCCATTGATCTGGATGACTTCACTGGCACTTTCTGCAACACAGGCCAAGTTTCCCTTAATC	1243
Qy	1141	TCTACTTTGAACAAAGCCCTTGGCATATCCCACTGAAGTTTGCACAGCTCCTGAGTGCCT	1200
Db	1244	TCCACCTTGAAGAGCCCTCGGCCCTGCAGATGCAAGTTTGCACGGCTCCAGCTCAGCC	1303
Qy	1201	TCGAGGCCAGTACTACTCTCC-----AGGAAGTGGAGTGGGGTGGAAAGTCC	1251
Db	1304	ATTGAGCAATAACTGCTGTCTCCAGTGGCAGGGGAACAGGAGGAGTAGCAGTCT	1363

```

Qy 1352 GGAGGAGCTCTGGAGGCGATGGATTCTGTGGCCGACAAAGAGCATGGCCCTCTACCCCTGTG 1311
Db 1364 GGAGGCGAGCTCGGGAGGCGATGCTGTGCTGTCTAGAGCCAAAGGCCCTCTACCCCGTG 1423
Qy 1312 GCAGTCACAGAAATGCTTTTGGCACTGCATCAATGAATCACATACCAGCAGCATTTGT 1371
Db 1424 GCAATTAACAGAAATGCTTCTGGCACTGGGTGTAATGGATGCAGTCAGTCACGACGAGACTGC 1483
Qy 1372 CAAGCAGGCGCTGTTTTTGATACCAAGCTGTAATTTGCTGCAACTGGCCCTGAACCTAA 1428
Db 1484 CAGCGCGGCTGTGCTTCGACACCAGCTGTGATTGCTGCAACTGGGCATAAACCTGA 1540

RESULT 3
US-10-202-436A-14
; Sequence 14, Application US/10202436A
; Publication No. US20030049261A1
; GENERAL INFORMATION:
; APPLICANT: ELIAS, JACK A
; APPLICANT: ZHU, ZHOU
; TITLE OF INVENTION: METHODS COMPOSITIONS AND KITS RELATING TO CHITINASES AND
; TITLE OF INVENTION: MOLECULES AND INFLAMMATORY DISEASE
; FILE REFERENCE: 044574-5107
; CURRENT APPLICATION NUMBER: US/10/202,436A
; CURRENT FILING DATE: 2002-10-17
; PRIOR APPLICATION NUMBER: U.S. 60/307,432
; PRIOR FILING DATE: 2001-07-24
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 1038
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: AMCase sense in situ hybridization probe
US-10-202-436A-14

```

	Query Match	28.2%;	Score 430;	DB 9;	Length 1038;
	Best Local Similarity	76.5%;	Pred. No. 7.8e-132;		
	Matches	598;	Conservative	0;	Mismatches 165; Indels. 19; Gaps 5;
QY	666	TGGGAGAAATAGTCTCTTTTACAATAACCCCTACTCGAGCTGGTAGCAATGCCTACCTCAA	725		
DB	1		60		
QY	726	TGTGGATTATGTCATGAACATATTGGAAGAACAATGGAGCCCACGCTGAGAAGCCTCATTTGT	785		
DB	61	TGTGGATTATGTCATGAACATATTGGAAGAACAATGGAGGCCACCCTGNGAAGCTCATCGT	120		
QY	786	TGGATTCGCCAGAGTATGACACACCTTCATCCTTGAGAACCCCCCTCTGATTAATGGAATTGG	845		
DB	121	TGGATTCCTTACCTATGACACAACCTTCATCTCGAGCAACCCCTCCAACACTTGGAAATTGG	180		
QY	846	TGCCCTTACCTCTGGTGATGGCCCTGCTGGGCCCTTATACACAGACAGGCTGGGTTCTGGGC	905		
DB	181	TGCCCCACACCTCTGGTGCTGCTGCTGGGCCCTATGCCAAGGAGTCTTGGAGATCTGGGC	240		
QY	906	CTACTATCAGATTTGCACCTTTCGAGAAGTGGAGCCACTCAGGCTCTGGGATGCTCCCA	965		
DB	241	TTACTACGAGATCTGTACCTTCTCGAAAAATGGAGCCACTCAGGAGTGGATGCCCCCTCA	300		
QY	966	AGAAGTGGCCTATGCCTATAAGGCCAACGAGTGGCTTGGCTTATGACAAATATCAAGAGCTT	1025		
DB	301	GGAAGTGGCTTATGCCTATCAGGCGAATGCTGGGTTGGCTTATGCAACGCTCAAGAGCTT	360		
QY	1026	CAGTGTTAAGGCTCAGTGGCTTAAGCAGAACAAATTTTGGAGGTGCCATGATCTGGGCCAT	1085		
DB	361	CGATATTAAGGCTCAAATGGCTTAAGCACAAACAATCTGGAGGCGCCATGGTCTGGGCCAT	420		
QY	1086	TGACCTTGATGACCTTCACTGGCTCTTCTGTGATCAGGGAANAATTTCTCTGACTTCTAC	1145		
DB	421	TGATCTGATGACTTCACTGGCACTTCTGCAACCGGCGAAGTTTTCCCTCTAACTCTCAC	480		

QY 1146 TTGAAACAAAGCCCTTGGCATATCCACTGAAGTTGCACAGCTCCTGACGTGCCCTCCGA 1205
DB 481 CCTGAAGAGCCCTCGGGCTGCAGAGTGAAGTTGCACAGCTCAGGCCATCAGCCATTGA 540
QY 1206 GCCAGTACACTCTCTCC-----AGGAAGTGGAGTGGGGTGGAGTCCGGAGG 1256
DB 541 GCCAATAACTGCTCTCCAGTGGCAGCGGGAACGGGAGGGAGTAGCAGCTCTGGAGG 600
QY 1257 AAGCTCTGGAGGAGTGGATCTCTGTGCC--GACAAACAGATGGCTCTACCC-----TG 1309
DB 601 CAGCTCGGAGGAGTGGATCTCTGTGGCAGAGCAACAGAGCTCTAACCCGTGGG 660
QY 1310 TGGCAGATGACAGAAATGCTTTTGGCAGTGCATCAATGGAATCACATACAGCAGCA-- 1367
DB 661 CAAATACCAGAGATGCTTCTGGGCACTGCGTGAATGAGTCAAGTACAGGAGCAGAAC 720
QY 1368 TTGTCAGCAGGAGGCTTCT--TTTGCATACAGCTGTAAATGCTGCAACTGGCCATGAACCT 1426
DB 721 TTGCCAGGCGGGCTGTCTCTCGAGCAGCTGTGAATGCTGCAACTGGGCAATTAACCT 780
QY 1427 AA 1428
DB 781 GA 782

RESULT 4

US-09-822-830A-402
; Sequence 402, Application US/09822830A
; Patent No. US20020142952A1
; GENERAL INFORMATION:
; APPLICANT: Genetics Institute, Inc.
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fechtel, Kim
; APPLICANT: Agostino, Michael J.
; APPLICANT: Howes, Steven H.
; APPLICANT: Resnick, Richard J.
; APPLICANT: Gulukota, Kamalakar
; APPLICANT: Graham, James R.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
; FILE REFERENCE: GIN 6402
; CURRENT APPLICATION NUMBER: US/09/822,830A
; CURRENT FILING DATE: 2001-03-29
; PRIOR FILING DATE: 2000-03-29
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 631
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 402
; LENGTH: 1391
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1353
; OTHER INFORMATION: n-a,c,g, or t
US-09-822-830A-402

Query Match 21.5%; Score 327.6; DB 10; Length 1391;
Best Local Similarity 58.1%; Pred. No. 1.1e-97;
Matches 667; Conservative 0; Mismatches 454; Indels 27; Gaps 4;
QY 23 CAGGCTGGCTCTTCTGCTGAATGCTCAGCTGGGGTCTGCCTACAATCTGATATGCTATT 82
DB 35 CAGGCTAGTGGTCTTCTGCTGCTCTCCAGGAGGATCTGCCTACAACTGGTTTGGGACT 94
QY 83 TCACCAACTGGGCGCCAGTATCGCCAGGCTTGGGGAGCTCAAGCTGATGACATTAAAC 142
DB 95 TTRCCAACTGGTCCCGAGGAGCCGAGGAACAGGAAATTCACCCCTGAGAATATTGACC 154
QY 143 CTGCGCTGTGTACTACCTGATCTATGCTTTCTGGGATGCGAAGCAATGAGATCACCA 202
DB 155 CCTTCTATGCTCTCATCTCATCTATTCATTCGCCAGCATCGAAGAACACAGGTTATCA 214

QY 203 CCATAGAAATGAATGATGTTTACTCTCTATAAAGCTTTCAATGACTTGAAGAAACAGGAACA 262
DB 215 TCAAGGACAAAGAGTGAAGTGTATGCTCTTACCAGACCATCAACAGTCTCAAAACCAAGAATC 274
QY 263 GCAAACTGAAACCCCTCTCGCAATTTGGAGGCTTGGAACTTTTGAAGTCTCTTCTTCACTA 322
DB 275 CCAAACTGAAATTTCTCTTCTCCATTTGGAGGTAACCTGTTTGGTTTCCAAAGGGTTCCACC 334
QY 323 CCATGTTTTCCTTCTCTCAACCGCCAGACCTTCAATTAACCTCAGTCAATCAATTTCTGC 382
DB 335 CTATGGTGGATCTCTTACATACGCTTGGAAATCAATTAACCTCAATTAACCTCTGTTCTGA 394
QY 383 GTCAATGTTGGTGTGATGAGTCTGACCTGACCTGGGAAATACCCAGGCTCAGTGGGAGCC 442
DB 395 GGAACCAATACTTTGATGGACTGGATGAAGTGGATCTACCCAGA----- 440
QY 443 CTCTCAGGACAAAGCATCTCTTCACTGTCTCTGGTGAAGAAATGCGTGAAGCTTTTGAGC 502
DB 441 -TCAGAAAGAAACACACTTCTTCACTGTCTGTGATTCATGAGTTAGCAGAAAGCCCTTTTCAGA 499
QY 503 AGGAGGCTATTGAGAGCAACAGCCAGACTGATGTTTACTGCTGTAGTCTGCTGGTGGGA 562
DB 500 AGGACTTCACAAATCCACCAAGAAAGGCTTCTTGTAGCTGGGGGTATCTGCAGGGA 559
QY 563 TTTCACATCCAGGCTGGCTATGAGATCCCTGAACCTTTCTAAAGTACCTGGATTTTCATCC 622
DB 560 GGCAATGATTTGATAACAGCTATCAAGTTGAGAAACTGGCAAAAGATCTGGATTTTCATCA 619
QY 623 ATGTCATGACATATGACCTCCATGGCTCTCGGGA-----GGGCTACATCTGGGAGATA 676
DB 620 ACCTCTGCTCTTGTGACTTCTTCCATGGGCTTTGGGAAAGGCCCTTATCACTGGCCACAACA 679
QY 677 GTCTCTTTTACAAATACCTTACTGAGACTGTGAGTGTAGTCAATCTCAATGTGGATTATG 736
DB 680 GCCCTCTGACCAAGGGTGGCAGGACAGAGGCGCAAGCTCTTACTACAAATGTGAATATG 739
QY 737 TCATGAACATATTGGAAGAACAAATGAGCCCCAGCTGAGAGCTCAATTTGGATTTCCAG 796
DB 740 CTGTGGGTACTGGATACATAAGGAATGCCATCAGAGAAGTGGTCAATGGGCATCCCA 799
QY 797 AGTATGGACACACCTTCACTCTGAGAAACCCCTCTGATATGAAATTTGGTGGCCCTACCT 856
DB 800 CATATGGGCACCTCTTCACTG-----GCCCTCTGCAAAAGCCCGTGGGGGGGGGGCTGCT 856
QY 857 CTGCTGATGGCCCTGCTGGGGCTTATACAGACAGGCTGGGTCTTGGGCTTACTATGAGA 916
DB 857 CTGGCCCTGGAGCTGCTGGACCATCACAGAGTCTTCAAGGCTTCTGGCTTATATGAGA 916
QY 917 TTTGACCTTTCTGAGAAAGTGGAGCCACTGAGGTCTGGGATGCGCTCCCAAGAGTGGCCT 976
DB 917 TCTGCCAGTCTCTGAAA--GGAGCCAAAGATCACAGAGCTCCAGGATCACAGGTTCCCT 973
QY 977 ATGCCCTATAGGCCAAAGTGGCTTGGCTATGACAAATACAAAGCTTCAGTGTAAAGG 1036
DB 974 ACAGCTCAGGGGACCCGGTGGGTGGCTATGATGATGAAAGATTTGGGGGCCCAAGG 1033
QY 1037 CTCAGTGGCTTAAAGCAGAAACAATTTTGGAGGTGCCATGATCTGGGCCATTTGACCTTGATG 1096
DB 1034 TTCAGTCTTAAAGATTTTAAACCTTGGGGTGGCTTGTATTTGGTCTTTTGACATGTTG 1093
QY 1097 ACTTCACTGGCTCTTCTGTGATCAGGGAATAATTTCTCTGACTTCTTCTTCTTGAACAAG 1156
DB 1094 ACTTCACTGGCAATCTGCAACCGGGGGCTTCTCCCTCTTGTGTCGAAGCAGTCAAGAGA 1153
QY 1157 CCCTTGGC 1164
DB 1154 GCCTTGGC 1161

RESULT 5

US-09-960-352-678
; Sequence 678, Application US/09960352
; Patent No. US2002013139A1


```

; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 678
; LENGTH: 449
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 03-LIB34-081-Q1-E1-A11
US-09-960-352-678

Query Match      21.5%; Score 327.2; DB 10; Length 449;
Best Local Similarity 84.4%; Pred. No. 6.8e-98;
Matches 368; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 30 GCGTCTTCGCTGAATGCTCAGCTGGGGTCTGCCTACATCTGATATGCTATTTACCAA 89
DB 14 GCGGCTCTGCGGAATGCTCAGCTAGGTTCTGCCTACCACTGGTATGCTACTTCTCTAA 73

QY 90 CTGGGCCAGTATCGGCCAGGCTCTGGGAGCTTCAAGCCTGATGACATTAACCCCTGCCT 149
DB 74 CTGGGCCAGTATCGGCCAGGCTCTGGGAGCTTCAAGCCTGATGACATTAACCCCTGCCT 133

QY 150 GTGTACTACCTGATGATGCTTGTCTGGGATGAGCAAAATGAGATCACCACCATAGA 209
DB 134 CTGCACTACCTGATGATGCTTGTCTGGGATGAGCAAAATGAGATCACCACCATAGA 193

QY 210 ATGGATGATGTTACTCTCTATTAAGCTTTCAATGACTTTGAAAACAGGAACAGAACT 269
DB 194 ATGGATGATGTTGCTCTCTATAGTCTTCAATGACTTTGAAAACAGGAACAGGAACAG 253

QY 270 GAAACCCCTCTGCGCAATGGAGGCTTGAACCTTTGGAACCTGCTTTCACCTACCATGT 329
DB 254 GAAATTCCTTGGCCATTTGGAGGCTTGAACCTTTGGAACCTGCTTTCACCTACCATGT 313

QY 330 TTCCACTTCTAGACCGCCAGCCTTCAATACCTGATGATCAAAATTTCTGGTCACTA 389
DB 314 TGCCACTTCCGAGACCGCAAGCTTTCATTTCTTCAGTCATCAAAATTTCTACACAGTA 373

QY 390 TGGGTTGATGACTGGACCTGGACTGGGAATACCCAGCTCAGCTGGGAGCCCTCTCA 449
DB 374 TGGATTGACGGGCTGGATTTTGTATGGGAGTACCCCTGGCTCTCGTGGAGCCCTCTCA 433

QY 450 GGACAAGCATCTCTTC 465
DB 434 GGACAAGCATCTCTTC 449

RESULT 6
US-10-097-340-44
; Sequence 44, Application US/10097340
; Publication No. US20030087250A1
; GENERAL INFORMATION:
; APPLICANT: John MONAHAN
; APPLICANT: Manjula GANNAVAPU
; APPLICANT: Sebastian HOERSCH
; APPLICANT: Shubhangi KAMATKAR
; APPLICANT: Steve G. KOVATS
; APPLICANT: Rachel E. MEYERS
; APPLICANT: Michael MORRISSEY
; APPLICANT: Peter OLANDT
; APPLICANT: Ami SEN
; APPLICANT: Peter VEIBY
; APPLICANT: Gordon B. MILLS
; APPLICANT: Robert C. BAST, JR.
; APPLICANT: Karen LU

```

```

; APPLICANT: Rosemarie SCHMANDT
; APPLICANT: Xumei ZHAO
; APPLICANT: Karen GLATT
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
; FILE REFERENCE: MRI-030
; CURRENT APPLICATION NUMBER: US/10/097,340
; CURRENT FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 60/276,025
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/325,149
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/276,026
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/324,967
; PRIOR FILING DATE: 2001/09/26
; PRIOR APPLICATION NUMBER: 60/311,732
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/325,102
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/323,580
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 563
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 1925
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-097-340-44

Query Match      21.0%; Score 319.6; DB 9; Length 1925;
Best Local Similarity 58.4%; Pred. No. 6.2e-95;
Matches 644; Conservative 0; Mismatches 434; Indels 24; Gaps 4;

QY 22 ACAGGCTCGGCTCTTCTGCTGAATGCTCAGCTGGGCTCTGCCTACAACTGATGATGCTAT 81
DB 148 ACAGGCTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 207

QY 82 TTCACCAACTGGCCCAAGTATCGCCAGCTCTGGGAGCTTCAAGCCTGATGACATTAAC 141
DB 208 TACACCAAGCTGTCCCGATGATGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 267

QY 142 CCCTGCTCTGTAATCACTCCTGATCTATGCTTGTGGGATGAGCAAAATGAGATCACC 201
DB 268 CGCTCTCTCTGTACCCACATCATCTACAGCTTTGCCAATATTAAGCAACGATCATCATCG 327

QY 202 ACATAGATGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 261
DB 328 ACCTGGGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 387

QY 262 AGCAAACTGAAACCTCTCTGCAATTTGGAGGCTGGAACCTTTGGAACCTGCTCTCTTCACT 321
DB 388 CCAACCTGAAGACTCTCTGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 447

QY 322 ACCATGTTTCCACTTCTCAGAACCGCCAGACCTTTCATTAACCTCAGTCAATCAAAATTTCTG 381
DB 448 AAGATAGCCTCCAAACACCCAGAGCTGCGCGGACTTTTCATCAAGTCACTACCGCATTTCTG 507

QY 382 CGTCAATGATGGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 441
DB 508 CGCACCCATGGCTTTGTGATGGGCTGGACCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 554

QY 442 CCTCTCTCAGCAAGCATCTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 501
DB 555 ---ACGAGAGAGCAAAACAGCATTTTACACCCCTAATCAAGGAATTAAGGCCGAATTTATA 612

QY 502 CAGGAGGCTATTGAGAGCAACAGCCAGACGATGATGATGATGATGATGATGATGATGATGATG 561
DB 613 AAGGAGGCCAGCCAGGGAAGAGC---AGCTCTCTCAGCGCAGCAGCTGCTGCTGCGGG 669

QY 562 ATTTCCAACTCCAGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 621
DB 670 AAGGTCAACCATGACAGCAGCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 729

```


Qy	231	TAAAGCTTTTCAA-TGACTTGGAAAACACAGGAACACAACTGAAAAACCTCTCTGCGCAATTG	289
Db	307	CCAGACCATCAACAGTTCTCAAAACCAAGATCCCAAACCTGAAAATTTCTCTGTGCCCAT	366
Qy	290	GAGGCTGGAACTTTGGAACTGCTCTCTTTCTACTACCATGTTTCCACTTCTCTCAGAACCGCC	349
Db	367	GAGGGTACCTTGTGTGGTTTCCAAAGGGTTCACCCCTATGTTGGGATCTTCTTACATPCAGCCT	426
Qy	350	AGACCTTTCAATACCTCAGTCAAAATTTCTGGGTCAGTATGGTTTGTATGGACTGGAC	409
Db	427	TGGAATTCATTAACCTCCATTAATCCTGTTCTGAGGAACCATAACTTTGTATGGACTGGATG	486
Qy	410	TGGACTGGGAATACCCAGGCTCAGCTGGGAGCCCTCTCAGGACAAGCATCTCTTCTACTG	469
Db	487	TAAGCTGGATCTACCCAGA-----TCAGAAAAGAAACACTCATTTTCACTG	531
Qy	470	TCCTGGTGAAGGAAATCGCTGAAGCTTTTCAGCAGGAGGCTATTTGAGAGCAACAGGCCA	529
Db	532	TGCTGATTCATGAGTTAGCAGAGACCTTTTCAGAAAGGACTTTCACAAAATCCCAACGAGAA	591
Qy	530	GACTGATGTTACTGTC-TGCTGTAGCTGGTGGGATTTTCCAACATCCAGGCTGGCTATGAG	588
Db	592	GGCTTCTCTTGACTCGCGGGGTATCTGCAGGGAGGCCAAATGATTGATACACAGCTATCAA	651
Qy	589	ATCCCTGAACCTTCTTAAGTACCTGGGATTTTCATCCATGTCATGACATATGACCTCCATGGC	648
Db	652	GTTTGAGAAACTGGCAAAAGATCTGGATTTTCATCAACCTCTGCTCTTGACTTCCATGGG	711
Qy	649	TCCTGGGA-----GGGCTACACTGGGGAAGATGCTCT-CTTTTACAAATACCTTACTGA	701
Db	712	TCTTGGGAAAAGCCCTTATCACTGGCCACAACAGCCCTGCTGAGCAAGGGGTGGCAGGA	771
Qy	702	GACTGTPAGCAATGCCTACCTCAATGTGGATTTATGTCAATGAACATATTGGGAAGACAATGG	761
Db	772	CAGAGGCCAAGCTCTCTACTACAATGTGGAATATGCTGTGGGTACTGGATACATAAGGG	831
Qy	762	AGCCCCAGCTGAGAGCTCAATTGTTGATTTCCAGAGATATGGACACACCTTTCATCTGAG	821
Db	832	AATGCCATCAGAGAAGGTGTCATGGGCATCCCCACATATGGGCA--CTCTTTCACACT	889
Qy	822	AAACCCCTCTGATAATGGAATTCGTGGCCCTACTCTGTGGTATGGCCCTGCTGGCGCTA	881
Db	890	GGCCTCTGCAGAAACCAACCGTGGGGCCCTGCTCTGGCCCTGGAGCTGCTGGACCCAT	949
Qy	882	TACCAGACAGGCTGGGTCTTGGGCCCTACTATGAGATTTGCACCTTTCTGAGAAGTGGAGC	941
Db	950	CACAGAGCTCTTCAGGCTTCTCGCCCTATTATGAGATCTGCCAGTTCCTGAAA---GGAGC	1006
Qy	942	CACTGAGGCTGGATGTCCTCCCAAGATGGCCCTATGCTCTATAGGCCAACAGTGGCT	1001
Db	1007	CAAGATCACGCGCTCCAGGATCAGCAGGTTCCCTACGCACTAAGGGGAACAGTGGGT	1066
Qy	1002	TGCTATGACAATATCAAGAGCTTCAGTGTTAAGGCTCAGTGGCTTAAGCAGAACATTT	1061
Db	1067	GGGCTATGATGATGTGAAGAGTATGGAGACCAAGGTTCACTTCTTAAGAAATTTAAACCT	1126
Qy	1062	TGGAGGTGCCATGATCTGGGCCATTAACCTTTGATGACTTCACTGGCTTCTTCTGTGATCA	1121
Db	1127	GGGAGGAGCCATGATCTGGTCTATTGACATGGATGACTTCACGTGGCAATCTCTGCACACCA	1186
Qy	1122	GGGAAATTTTCTCTGACTTCTACTTTGAAACAAGCCCTTGGC	1164
Db	1187	GGGCCCTTACCTCTCTGTCCAAAGCAGTCAAGAGAAGCCTTGGC	1229

RESULT 10

US-09-960-352-2589	US/09960352
; Sequence 2589, Application	
; Patent No. US20020137139A1	
; GENERAL INFORMATION:	
; APPLICANT: Warren, Wesley C.	
; APPLICANT: Tao, Nengbing	
; APPLICANT: Byatt, John C.	

```

; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 2589
; LENGTH: 415
; TYPE: DNA
; ORGANISM: Bos taurus.
; OTHER INFORMATION: Clone ID: 12-LIB188-006-Q1-E1-C7
; US-09-960-352-2589

```

Query Match 10.0%; Score 152.2; DB 10; Length 415;
Best Local Similarity 63.6%; Pred. NO. 1e-39;
Matches 232; Conservative 0; Mismatches 133; Indels 0; Gaps 0

84	QY	CACCAACTGGGCCCAGTATCGGCCAGAGTCTCGGGAGCTTCAAGCCTGTATGATCAATTAACCC	143
12	Db	CGCCATGGCCCGGGATACCGGGGAAGTGATGGGAGCTGCTTCCACAGACGCCATGCAGCC	71
144	QY	CTGCCTGTGTACTCACCTGATCTATGCCCTTGTCTGGGATCGAGAACATGAGATCACCAC	203
72	Db	CTTCCTGTGCACCCATGTCATCTACAGCTTTGCCAACATAGCAACAATGAGATCGACAC	131
204	QY	CATAGAATGGAATGATGTTACTCTCTATAAAGCTTTCAATGACTTTGAAAAACAGGAACAG	263
132	Db	CTGGGAGTGGNAATGACGTGACGCTCTATGACACACTGAACACTCAAGAACAGGAACCC	191
264	QY	CAAACTGAAACCCCTCCTGGCAAATGGAGGTGGAACTTTGGAAGTGTCTCTTTTCACTAC	323
192	Db	CAACCTGAAGACCCCTCCTATCTGTGGAGGATGGAACCTCGGTTCTCAAGAAGATTTTCCAA	251
324	QY	CATGGTTTCCACTTCTCAGAACCCGCAGACCTTCATTACCTCAGTCAATCAAAATTTCTGGC	383
252	Db	GATAGCTTCCAAAGACCCAGAGTCGCGAGGACTTTCATCAAGTCGGTGCACCACTTTCTGGC	311
384	QY	TCAGTATGGGTTTCATGGACTGGACCTGGACTGGGAATACCCAGGCTCACGTGGGAGCC	443
312	Db	GACCCATGGCTTTCATGGACTGGACCTAGCTCTACCCCGGTTGGAGAGACAGCG	371
444	QY	TCCTC 448	
372	Db	GCATC 376	

RESULT 11

```

RESULT II
US-09-960-352-13050
; Sequence 13050, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warten, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND O
; TITLE OF INVENTION: MUSCLE AND FAT DE
; FILE REFERENCE: 16511.006/37-21(102398)
; CURRENT APPLICATION NUMBER: US/09/960,
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 13050
; LENGTH: 414
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 56-L1B18
US-09-960-352-13050

```

Query Match 9.8%; Score 149.6; DB 10; Length 414;
Best Local Similarity 64.4%; Pred. NO. 7.4e-39;
Matches 224; Conservative 0; Mismatches 124; Indels 0; Gaps 0

RESULT 10
US-09-960-352-2589

QY 101 ATCGGCCAGGCTGCGGAGCTTCAAGCCTGATGACATTAACCCCGCCTGTGTACTCACC 160
Db 1 ACCGGGAGGCTGATGGAGCTGCTTCCAGAGCCCATCGACCCCTTCTGTGCACCCATG 60
QY 161 TGATCTATGCTTCTGCTGGGATGAGAACAAATGAGATCACCACCATAGAAATGAATGATG 220
Db 61 TCATCTACAGCTTTGCCAACATAAGCAACAATGAGATGACACCTGGGAGTGAATGAGC 120
QY 221 TTACTCTCTATAAAGCTTCAATGACTTGAAGAACAGGAAACAGCAAACTGAAACCCCTCC 280
Db 121 TGACCTCTATGACACATGAACACATCAAGAACAGGAAACCCCAACCTGGAAGCCCTCC 180
QY 281 TGGCAATGAGGCTTGAACCTTTGGAAGCTGCTCTTCACTACCATGTTTCCACTTCTC 340
Db 181 TATCTGTTGGAGATGGAACCTTCGCTTCTCAAAGATTTTCCAAGATAGCTTCCAAGACC 240
QY 341 AGAACCGCCAGACCTTCAATACCTCAGTCATCAAAATTTCTGGTGCAGTATGGGTTGATG 400
Db 241 GGAGTCGCGAGGACTTTCATCAAGTCGGTGCCACCAATTTCTGCGGACCCATGCTTTGATG 300
QY 401 GACTGGACCTGGACTGGGAATACCCAGGCTCAGTGGGAGCCCTCCTC 448
Db 301 GACTGGACCTAGCATGGCTCTACCCCGGGTGGAGAGACAGCGGCATC 348

RESULT 12
US-09-960-352-7364
; Sequence 7364, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 7364
; LENGTH: 410
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 32-LIB188-023-Q1-E1-H11
US-09-960-352-7364

Query Match 9.7%; Score 148; DB 10; Length 410;
Best Local Similarity 64.1%; Pred. No. 2.5e-38;
Matches 223; Conservative 0; Mismatches 125; Indels 0; Gaps 0;
QY 101 ATCGGCCAGGCTGCGGAGCTTCAAGCCTGATGACATTAACCCCGCCTGTGTACTCACC 160
Db 1 ACCGGGAGGCTGATGGAGCTGCTTCCAGAGCCCATCGACCCCTTCTGTGCACCCATG 60
QY 161 TGATCTATGCTTCTGCTGGGATGAGAACAAATGAGATCACCACCATAGAAATGAATGATG 220
Db 61 TCATCTACAGCTTTGCCAACATAAGCAACAATGAGATGACACCTGGGAGTGAATGAGC 120
QY 221 TTACTCTCTATAAAGCTTCAATGACTTGAAGAACAGGAAACAGCAAACTGAAACCCCTCC 280
Db 121 TGACCTCTATGACACATGAACACATCAAGAACAGGAAACCCCAACCTGGAAGCCCTCC 180
QY 281 TGGCAATGAGGCTTGAACCTTTGGAAGCTGCTCTTCACTACCATGTTTCCACTTCTC 340
Db 181 TATCTGTTGGAGATGGAACCTTCGCTTCTCAAAGATTTTCCAAGATAGCTTCCAAGACC 240
QY 341 AGAACCGCCAGACCTTCAATACCTCAGTCATCAAAATTTCTGGTGCAGTATGGGTTGATG 400
Db 241 GGAGTCGCGAGGACTTTCATCAAGTCGGTGCCACCAATTTCTGCGGACCCATGCTTTGATG 300
QY 401 GACTGGACCTGGACTGGGAATACCCAGGCTCAGTGGGAGCCCTCCTC 448
Db 301 GACTGGACCTAGCATGGCTCTACCCCGGGTGGAGAGACAGCGGCATC 348

Db 301 GACTGGACCTAGCATGGCTCTACCCCGGGTGGAGAGACAGCGGCATC 348
RESULT 13
US-09-960-352-3057
; Sequence 3057, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 3057
; LENGTH: 424
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 14-LIB188-002-Q1-E1-D5
US-09-960-352-3057

Query Match 9.7%; Score 148; DB 10; Length 424;
Best Local Similarity 64.1%; Pred. No. 2.6e-38;
Matches 223; Conservative 0; Mismatches 125; Indels 0; Gaps 0;
QY 101 ATCGGCCAGGCTGCGGAGCTTCAAGCCTGATGACATTAACCCCGCCTGTGTACTCACC 160
Db 1 ACCGGGAGGCTGATGGAGCTGCTTCCAGAGCCCATCGACCCCTTCTGTGCACCCATG 60
QY 161 TCATCTATGCTTCTGCTGGGATGAGAACAAATGAGATCACCACCATAGAAATGAATGATG 220
Db 61 TCATCTACAGCTTTGCCAACATAAGCAACAATGAGATCAGACCTGGGAGTGAATGATG 120
QY 221 TTACTCTCTATAAAGCTTCAATGACTTGAAGAACAGGAAACAGCAAACTGAAACCCCTCC 280
Db 121 TGACCTCTATGACACATGAACACATCAAGAACAGGAAACCCCAACCTGGAAGCCCTCC 180
QY 281 TGGCAATGAGGCTTGAACCTTTGGAAGCTGCTCTTCACTACCATGTTTCCACTTCTC 340
Db 181 TATCTGTTGGAGATGGAACCTTCGCTTCTCAAAGATTTTCCAAGATAGCTTCCAAGACC 240
QY 341 AGAACCGCCAGACCTTCAATACCTCAGTCATCAAAATTTCTGGTGCAGTATGGGTTGATG 400
Db 241 GGAGTCGCGAGGACTTTCATCAAGTCGGTGCCACCAATTTCTGCGGACCCATGCTTTGATG 300
QY 401 GACTGGACCTGGACTGGGAATACCCAGGCTCAGTGGGAGCCCTCCTC 448
Db 301 GACTGGACCTAGCATGGCTCTACCCCGGGTGGAGAGACAGCGGCATC 348

RESULT 14
US-09-960-352-3072
; Sequence 3072, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 3072
; LENGTH: 418
; TYPE: DNA
; ORGANISM: Bos taurus

OTHER INFORMATION: Clone ID: 14-LIB188-015-Q1-E1-D5
US-09-960-352-3072

Query Match 9.7%; Score 147.4; DB 10; Length 418;
Best Local Similarity 63.9%; Pred. No. 4e-38;
Matches 223; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

QY 100 TATGGCCAGGTCTGGGAGCTTCAAGCCTGATGACATTAAACCCCTGCTGTGTACTCAC 159
DB 11 TACCGGAGGATGATGGAGCTGCGTCCAGAGCCCATCGACCTTCCTGTGCACCCAT 70
QY 160 CTGATCATGCCCTTGTCTGGGATGCAGAACAAATGAGATCACCACCATAGAAATGAATGAT 219
DB 71 GTCATCGACAGCTTGGCAACATAAAGCAACAATGAGATGACACCTGGGAGTGAATGAC 130
QY 220 GTTACTCTCTATAAGCTTTCAATGACTTGAAGAGGAGACAGCAACACTGAAACCCCTC 279
DB 131 GTGAGGCTTATGACACTGAAACACTCAAGAACAGGAAACCCCAACCTGGAAGACCCCTC 190
QY 280 CTGGCAATTGGAGCTGGAACTTTGGAACCTGCTCTTCACTACCACTGGTTTCCACTTCT 339
DB 191 CTATCTGTTGAGAGTGGAACTTCGGTTCTCAAGATTTTCCAGATAGCTTCCAAGACT 250
QY 340 CAGAACGCCAGACCTTCATTACTCAGTCAATCAAAATTTCTGCTCAGTATGGGTTTGTAT 399
DB 251 CAGAGTGCAGAGCTTTTCATCAAGTGGTGCCACCAATTTCTGGGAGCCCATGGCTTTGAT 310
QY 400 GGACTGGACCTGGACTGGGAATACCCAGGCTCAGTGGGAGCCCTCCTC 448
DB 311 GGACTGGACCTAGCATGGCTCTACCCCGGTGGAGAGACAAGCGGCATC 359

RESULT 15

US-09-960-352-265/c

Sequence 265, Application US/09960352

Patent No. US20020137139A1

GENERAL INFORMATION:

APPLICANT: Warren, Wesley C.

APPLICANT: Tao, Nengbing

APPLICANT: Byatt, John C.

APPLICANT: Mathialagan, Nagappan

TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND

FILE OF INVENTION: MUSCLE AND FAT DEPOSITION

FILE REFERENCE: 16511.006/37-21(10298)C

CURRENT APPLICATION NUMBER: US/09/960,352

CURRENT FILING DATE: 2001-09-24

NUMBER OF SEQ ID NOS: 15112

SEQ ID NO 265

LENGTH: 399

TYPE: DNA

ORGANISM: Bos taurus

OTHER INFORMATION: Clone ID: 02-LIB188-020-Q1-E1-A5

US-09-960-352-265

Query Match 8.9%; Score 136.2; DB 10; Length 399;
Best Local Similarity 65.0%; Pred. No. 2.1e-34;
Matches 201; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 140 ACCCTGCTGTCTACTACCTGATCTATCCCTTCTGCTGGGATGCAGAACATGAGATCA 199
DB 399 ACCCTTCTGTGCCCATGTGCTATCAGCTTTGCCCAACATAGCAACAATGAGATCG 340
QY 200 CCACCATAGAATGAATGATGTACTCTCTATAAAGCTTCAATGACTTGAACAAACAGGA 259
DB 339 ACACCTGGGAGTGAATGAGTGCACGCTCTATCACACACTGAACACACTCAAGAACAGGA 280
QY 260 ACAGAACTGAACCCCTCTCGGCAATTGGAGGCTGGAACTTTGAACTGCTCCTTTCA 319
DB 279 ACCCAACCTGAGACCCCTCTATCTGTGGAGGATGGAACCTTCGGTTCTCAAAGATTTT 220
QY 320 CTACCATGTTTCCATCTTCAGACCCGACGACCTTCACTACCTCAGTCATCAATTTTC 379
DB 219 CCAGATAGCTTCCAAAGACCCAGAGTCGCGAGGACTTTTCATCAAGTCGTCGCCACCATTTTC 160

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run On: July 3, 2003, 04:32:16 ; Search time 2287.02 Seconds
(without alignments)
10799.277 Million cell updates/sec

Title: US-10-004-219B-3
Perfect score: 1525
Sequence: 1 atggccgaagtacttctcgt.....taaaattgtagcaaaaca 1525

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

- 1: em_estba:**
- 2: em_esthum:**
- 3: em_estin:**
- 4: em_estmu:**
- 5: em_estov:**
- 6: em_estpl:**
- 7: em_estro:**
- 8: em_hic:**
- 9: gb_est1:**
- 10: gb_est2:**
- 11: gb_hic:**
- 12: gb_est3:**
- 13: gb_est4:**
- 14: gb_est5:**
- 15: em_estfun:**
- 16: em_estom:**
- 17: gb_gss:**
- 18: em_gss_hum:**
- 19: em_gss_inv:**
- 20: em_gss_pln:**
- 21: em_gss_vrt:**
- 22: em_gss_fun:**
- 23: em_gss_mam:**
- 24: em_gss_mus:**
- 25: em_gss_Other:**
- 26: em_gss_pro:**
- 27: em_gss_rod:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length	Description
1	1512.2	99.2 1533 11	AK008633 Mus muscu
2	1510	99.0 1535 11	AK008757 Mus muscu
3	1488.6	97.6 1534 11	AK008650 Mus muscu
4	1465.8	96.1 1530 11	AK007573 Mus muscu
5	861.2	56.5 892 14	BQ231098 AGENCOURT
6	851.6	55.8 1007 9	AV072529 AV072529

7	829.4	54.4	872	14	BQ231064	BQ231064 AGENCOURT
8	823.6	54.0	838	12	BG867815	BG867815 602786336
9	801.4	52.6	837	12	BG868741	BG868741 602787570
10	801.2	52.5	855	12	BG872136	BG872136 602792974
11	796.2	52.2	836	12	BG869655	BG869655 602789716
12	794.6	52.1	880	12	BG871468	BG871468 602790678
13	789.8	51.8	894	12	BG869958	BG869958 602789826
14	788.4	51.7	847	12	BG866775	BG866775 602786505
15	788	51.7	809	12	BG866417	BG866417 602785352
16	785	51.5	833	12	BG871337	BG871337 602790509
17	784.6	51.4	863	12	BG872850	BG872850 602793973
18	783	51.3	874	12	BG865756	BG865756 602784345
19	776.6	50.9	856	12	BG870147	BG870147 602793390
20	775.8	50.9	948	13	BI664452	BI664452 603290028
21	775.4	50.8	845	12	BG865848	BG865848 602788264
22	774.8	50.8	804	12	BG872805	BG872805 602793925
23	774.6	50.8	839	12	BG870376	BG870376 602791269
24	774.6	50.8	845	12	BG867145	BG867145 602786759
25	774.2	50.8	822	12	BG870367	BG870367 602791260
26	772.4	50.6	837	12	BG865578	BG865578 602783726
27	771.2	50.6	843	12	BG868207	BG868207 602784727
28	771	50.6	1155	12	BG867473	BG867473 602788579
29	765.8	50.2	939	14	BQ219632	BQ219632 AGENCOURT
30	763.4	50.1	799	12	BG872144	BG872144 602792982
31	763.4	50.1	848	12	BG869143	BG869143 602789020
32	762.8	50.0	769	12	BG868463	BG868463 602785844
33	762.8	50.0	794	12	BG869492	BG869492 602789190
34	762.2	50.0	838	12	BG870392	BG870392 602791292
35	762	50.0	886	12	BG870840	BG870840 602791891
36	760	49.8	882	12	BG873488	BG873488 602791730
37	759.8	49.8	845	12	BG865076	BG865076 602784229
38	759.8	49.8	900	12	BG867646	BG867646 602787435
39	759.8	49.8	949	12	BG868019	BG868019 602788381
40	759.2	49.8	857	12	BG871434	BG871434 602790536
41	759	49.8	824	12	BG865859	BG865859 602786278
42	757.8	49.7	831	12	BG867708	BG867708 602786811
43	757.4	49.7	884	12	BG867239	BG867239 602786266
44	756.6	49.6	838	12	BG868956	BG868956 602784563
45	755.8	49.6	830	12	BG865815	BG865815 602788220

ALIGNMENTS

RESULT 1	AK008633	AK008633	1533 bp	mrna	linear	HTC 19-JAN-2002
LOCUS	AK008633	Mus musculus adult male stomach cDNA, RIKEN full-length enriched library, clone:2200003E03:Chitinase, acidic, full insert sequence.				
DEFINITION	AK008633	AK008633				
ACCESSION	AK008633	AK008633	1	GI:12842941		
VERSION	AK008633	HTC, CAP trapper.				
KEYWORDS	HTC, CAP trapper.	Mus musculus (strain:C57BL/6J) adult male stomach cDNA to mRNA, clone_lib:RIKEN full-length enriched mouse cDNA library				
SOURCE		clone:2200003E03.				
ORGANISM		Mus musculus				
REFERENCE		Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
AUTHORS		1				
TITLE		Carninci, P. and Hayashizaki, Y.				
JOURNAL		High-efficiency full-length cDNA cloning				
MEDLINE		Meth. Enzymol. 303, 19-44 (1999)				
PUBMED		99279253				
REFERENCE		2				
AUTHORS		Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.				
TITLE		Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
JOURNAL		Genome Res. 10 (10), 1617-1630 (2000)				
MEDLINE		20499374				
PUBMED		11042159				
REFERENCE		3				

Db	497	GAATGCGTGAAGCTTTTGAGCAGGAGGCTATTGAGAGCAACAGGCCACGACTGATGGTT	556
Qy	541	ACTGCTGCTAGCTGGTGGGATTTTCAACATCCAGGCTGGCTATGAGATCCCTGAACTT	600
Db	557	ACTGCTGCTAGCTGGTGGGATTTTCAACATCCAGGCTGGCTATGAGATCCCTGAACTT	616
Qy	601	TCTAAGTACCTGGATTCATCCATGTCATGACATATGACATCCATGGCTCCTGGGAGGC	660
Db	617	TCTAAGTACCTGGATTCATCCATGTCATGACATATGACATCCATGGCTCCTGGGAGGC	676
Qy	661	TACACTGGGAGAAATAGTCTCTTTACAAATACCTACTGAGACTGGTACGANTGCCTAC	720
Db	677	TACACTGGGAGAAATAGTCTCTTTACAAATACCTACTGAGACTGGTACGANTGCCTAC	736
Qy	721	CTCAATGTGATTATGTCAATGAACATTGGAAGAACAAATGGAGCCCGAGCTGAGAAGCTC	780
Db	737	CTCAATGTGATTATGTCAATGAACATTGGAAGAACAAATGGAGCCCGAGCTGAGAAGCTC	796
Qy	781	ATTGTTGGATTCCAGAGTATGGACACACCTTCATCCTGAGAAACCCCTCTGATAATGGA	840
Db	797	ATTGTTGGATTCCAGAGTATGGACACACCTTCATCCTGAGAAACCCCTCTGATAATGGA	856
Qy	841	ATTGGTGGCCCTACCTCTGTGATGGCCCTGCTGGGCCCTATACCAGACAGGCTGGGTT	900
Db	857	ATTGGTGGCCCTACCTCTGTGATGGCCCTGCTGGGCCCTATACCAGACAGGCTGGGTT	916
Qy	901	TGGGCCCTACTATGAGATTTCACGCTTCTCAGAAAGTGGAGCCACTGAGGTCCTGGGATGCC	960
Db	917	TGGGCCCTACTATGAGATTTCACGCTTCTCAGAAAGTGGAGCCACTGAGGTCCTGGGATGCC	976
Qy	961	TCCAGAAAGTGCCCTATAGCCTATAGAGCCAAAGAGTGGCTTGGCTATGACAATACAAG	1020
Db	977	TCCAGAAAGTGCCCTATAGCCTATAGAGCCAAAGAGTGGCTTGGCTATGACAATACAAG	1036
Qy	1021	AGCTTCAGTCTTAAGGCTCAGTGGCTTAAGCAAGCAAAATTTGAGAGTGCCATGACTGG	1080
Db	1037	AGCTTCAGTCTTAAGGCTCAGTGGCTTAAGCAAGCAAAATTTGAGAGTGCCATGACTGG	1096
Qy	1081	GCCATTGACCTTGATGACTTTCACCTGGGCTCTTTCTGTGATCAGGGAATAATTCCTCTGACT	1140
Db	1097	GCCATTGACCTTGATGACTTTCACCTGGGCTCTTTCTGTGATCAGGGAATAATTCCTCTGACT	1156
Qy	1141	TCTACTTTGAACAAAGCCCTTGGCATATCCACTGAGAGTTGCACAGCTCCTGAGCTGCC	1200
Db	1157	TCTACTTTGAACAAAGCCCTTGGCATATCCACTGAGAGTTGCACAGCTCCTGAGCTGCC	1216
Qy	1201	TCCGAGCCAGTACTACTCCTCCAGGAAGTGGAGTGGGGTGGAGCTCCGAGGAAGC	1260
Db	1217	TCCGAGCCAGTACTACTCCTCCAGGAAGTGGAGTGGGGTGGAGCTCCGAGGAAGC	1276
Qy	1261	TCTGGAGGCAGTGGATTCTGTGCGCAAGCAGATGGCCCTCTACCCCTGTGGCAGATGAC	1320
Db	1277	TCTGGAGGCAGTGGATTCTGTGCGCAAGCAGATGGCCCTCTACCCCTGTGGCAGATGAC	1336
Qy	1321	AGAAATGCTTTTGGCAGTGCCATCAATGGAATCACATACCAAGCAGANTTGTCAAGCAGG	1380
Db	1337	AGAAATGCTTTTGGCAGTGCCATCAATGGAATCACATACCAAGCAGANTTGTCAAGCAGG	1396
Qy	1381	CTTGTTTTTCATACCAAGCTGTAATGCTGCAACTGGCCATGAACCTTAATGCCATTCTTCC	1440
Db	1397	CTTGTTTTTCATACCAAGCTGTAATGCTGCAACTGGCCATGAACCTTAATGCCATTCTTCC	1456
Qy	1441	AGAAATTCCTGCACTCTCCCTTACTCCTCACCAAAAGTAATCTATCTTCCCTTTAACTTA	1500
Db	1457	AGAAATTCCTGCACTCTCCCTTACTCCTCACCAAAAGTAATCTATCTTCCCTTTAACTTA	1516
Qy	1501	TGCAATAAAATTTGGTAG	1517
Db	1517	TGCAATAAAATTTGGTTG	1533

Query Match		97.6%; Score 1488.6; DB 11; Length 1534;
Best Local Similarity		99.6%; Pred. No. 0;
Matches 1513; Conservative		0; Mismatches 4; Indels 2; Gaps 2;
QY	1	ATGCGCAAGTACTTCTCGTCAGAGTCTGGCTCTTCTGCTGAATGCTCAGCTGGGTCT 60
DB	17	ATGCGCAAGTACTTCTCGTCAGAGTCTGGCTCTTCTGCTGAATGCTCAGCTGGGTCT 76
QY	61	GCCTACAATCTGATATGCTATTTTCAACCACTGGGCCAGTATCGGCCAGTCTGGGGAGC 120
DB	77	GCCTACAATCTGATATGCTATTTTCAACCACTGGGCCAGTATCGGCCAGTCTGGGGAGC 136
QY	121	TTCAAGCTGATGACATTAACCCCTGCTGTGTACTACCTGATCATGCTTGTGTGG 180
DB	137	TTCAAGCTGATGACATTAACCCCTGCTGTGTACTACCTGATCATGCTTGTGTGG 196
QY	181	ATSCAGAACATGAGATCAACCACTAGATGAATGATTTACTCTCTATAAGCTTTC 240
DB	197	ATSCAGAACATGAGATCAACCACTAGATGAATGATTTACTCTCTATAAGCTTTC 256
QY	241	AATGACTTGAACCAAGGACAGCAAACTGAAACCCCTCTGGCAATTTGGAGCTGGAAC 300
DB	257	AATGACTTGAACCAAGGACAGCAAACTGAAACCCCTCTGGCAATTTGGAGCTGGAAC 316
QY	301	TTTGGAACTGCTCTTCTACCTACCATGTTTCCACTTCTCAGAACCCGACAGCTTCATT 360
DB	317	TTTGGAACTGCTCTTCTACCTACCATGTTTCCACTTCTCAGAACCCGACAGCTTCATT 376
QY	361	ACCTCAGTCATCAAAATTTCTGCTCAGTATGGTTTGTGAGCTGGACCTGGAGTGGAA 420
DB	377	ACCTCAGTCATCAAAATTTCTGCTCAGTATGGTTTGTGAGCTGGACCTGGAGTGGAA 436
QY	421	TACCCAGGCTCAGTGGGAGCCCTCTCAGGACAAGCATCTTCTACTGTCTGGTGAAG 480
DB	437	TACCCAGGCTCAGTGGGAGCCCTCTCAGGACAAGCATCTTCTACTGTCTGGTGAAG 496
QY	481	GAATCCGTGAAGCTTTTGAAGAGGAGCTATTGAGAGCAACAGGCCAGACTGATGGTT 540
DB	497	GAATCCGTGAAGCTTTTGAAGAGGAGCTATTGAGAGCAACAGGCCAGACTGATGGTT 556
QY	541	ACTGCTGCTGAGTGGTGGATTTCACATCCAGGCTGGCTATCAGATCCCTGAACTT 600
DB	557	ACTGCTGCTGAGTGGTGGATTTCACATCCAGGCTGGCTATCAGATCCCTGAACTT 616
QY	601	TCTAAGTACCTGGATTTTCACTGATGATGATGATGATGATGATGATGATGATGATGAT 660
DB	617	TCTAAGTACCTGGATTTTCACTGATGATGATGATGATGATGATGATGATGATGATGAT 676
QY	661	TACACTGGGAGATAGTCTCTTTTACAATACCTCTAGACTGGTGGTGGTGGTGGTGGT 720
DB	677	TACACTGGGAGATAGTCTCTTTTACAATACCTCTAGACTGGTGGTGGTGGTGGTGGT 735
QY	721	CTCAATGTGGATTTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
DB	736	CTCAATGTGGATTTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 795
QY	781	ATTGTTGGATTTCCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
DB	796	ATTGTTGGATTTCCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 854
QY	841	ATTGTTGGATTTCCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
DB	855	ATTGTTGGATTTCCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 914
QY	901	TGGGCTACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
DB	915	TGGGCTACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 974
QY	961	TCCCAAGAAGTCCCTATGCTTATAGGCCAAGAGTGGCTTGGCTTATGACATATCAAG 1020
DB	975	TCCCAAGAAGTCCCTATGCTTATAGGCCAAGAGTGGCTTGGCTTATGACATATCAAG 1034
QY	1021	AGCTTCAGTGTAAAGCTCAGTGGCTTATAGGAGACAAATTTTGGAGGTGCCATGATCTGG 1080

Db	1035	AGCTTCAGTGTAAAGCTCAGTGGCTTAAAGCAGAACAAATTTTGGAGGTGCCATGATCTGG 1094
QY	1081	GCCATTGACCTTGCATGACTTCACTGGCTCTTTCTGTGATCAGGAGAAATTTTCTCTGACT 1140
Db	1095	GCCATTGACCTTGCATGACTTCACTGGCTCTTTCTGTGATCAGGAGAAATTTTCTCTGACT 1154
QY	1141	TCTACTTTGAACAAAGCCCTTGGCATATCCACTGAAGGTTGACAGCTCCTGACGTGCCT 1200
Db	1155	TCTACTTTGAACAAAGCCCTTGGCATATCCACTGAAGGTTGACAGCTCCTGACGTGCCT 1214
QY	1201	TCCGAGCCAGTGAATCTCTCCAGGAAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGG 1260
Db	1215	TCCGAGCCAGTGAATCTCTCCAGGAAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGG 1274
QY	1261	TCTGGAGGAGTGAATCTCTCCAGGAAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGG 1320
Db	1275	TCTGGAGGAGTGAATCTCTCCAGGAAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGG 1334
QY	1321	AGAAATGCTTTTGGCAGTGCATCAATGGAATCACAATACAGCAGCAATTTGTCAGCAGGG 1380
Db	1335	AGAAATGCTTTTGGCAGTGCATCAATGGAATCACAATACAGCAGCAATTTGTCAGCAGGG 1394
QY	1381	CTTGTTTTGATACAGAGTGTAAATGCTGCACTGGCCATGAACCTTAATGCCATTTCTTCC 1440
Db	1395	CTTGTTTTGATACAGAGTGTAAATGCTGCACTGGCCATGAACCTTAATGCCATTTCTTCC 1454
QY	1441	AGAAATTTCTGCACTCTCTTCTACTCTCACCACCAAGTAACTATCTTCCCTTTAACTTA 1500
Db	1455	AGAAATTTCTGCACTCTCTTCTACTCTCACCACCAAGTAACTATCTTCCCTTTAACTTA 1514
QY	1501	TGCAATAAAATTTGGTAGCC 1519
Db	1515	TGCAATAAAATTTGGTAGCC 1533

RESULT 4
AK007573
LOCUS
DEFINITION
Mus musculus 10 day old male pancreas cDNA, RIKEN full-length enriched library, clone:1810021D23:chitinase, acidic, full insert sequence.
ACCESSION
AK007573
VERSION
AK007573.1 GI:12841201
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (strain:C57BL/6J) 10 day old male pancreas cDNA to mRNA, clone.lib:RIKEN full-length enriched mouse cDNA library
clone:1810021D23.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
PUBMED
10349636
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
PUBMED
11042159
3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsumai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

TITLE	RIXEN integrated sequence analysis (RISA) system - 384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE	20530913
PUBMED	11076851

AUTHORS

Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, M., Nishik, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fieischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaudo, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Boujunga, N., Carninci, P., de Bonaudo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyono, O., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohsaki, S., and Hayashizaki, Y.

TITLE
Functional annotation of a full-length mouse cDNA collection
JOURNAL
Nature 409 (1992), 685-690 (1992)

QY 841 ATTGGTCCCTACCTCTGGTGATGGCCCTGCTGGCCCTATACACAGACAGGCTGGGTTCC 900
 Db 856 --TGGTCCCTACCTCTGGTGAT--GCCCTGGTGGC--CCTATACACAGACAGGCTGGGTTCC 911
 QY 901 TGGCCCTACTATGAGATTGCACCTTTCTGAGAGTGGAGCCACTGAGGCTGGGATGCC 960
 Db 912 TGGCCCTACTATGAGATTGCACCTTTCTGAGAGTGGAGCCACTGAGGCTGGGATGCC 971
 QY 961 TCCCAAGAAGTGGCCCTATGCTATAGGCCCAACAGTGGCTTGGCTATGACATATCAAG 1020
 Db 972 TCCCAAGAAGTGGCCCTATGCTATAGGCCCAACAGTGGCTTGGCTATGACATATCAAG 1031
 QY 1021 AGCTTCAGTGTAAAGGCTAGTGGCTTAAAGCAGAACAAATTTGGAGTGGCCATGCTGG 1080
 Db 1032 AGCTTCAGTGTAAAGGCTAGTGGCTTAAAGCAGAACAAATTTGGAGTGGCCATGCTGG 1091
 QY 1081 GCCATTGACCTTGATGACTTCACTGGCTCTTTCTGTGATCAGGGAATTTCTCTGACT 1140
 Db 1092 GCCATTGACCTTGATGACTTCACTGGCTCTTTCTGTGATCAGGGAATTTCTCTGACT 1151
 QY 1141 TCTACTTTGAACAAAGCCCTTGGCAATCCACTGAAGTTTGCACAGCTCCTGACGTGCCT 1200
 Db 1152 TCTACTTTGAACAAAGCCCTTGGCAATCCACTGAAGTTTGCACAGCTCCTGACGTGCCT 1211
 QY 1201 TCCGAGCAGTACTACTCTCCAGGAAGTGGAGTGGGGTGGAGCTCCGGAGGAAGC 1260
 Db 1212 TCCGAGCAGTACTACTCTCCAGGAAGTGGAGTGGGGTGGAGCTCCGGAGGAAGC 1271
 QY 1261 TCTGGAGGAGTGGATCTGTGCGGACAAAGCAGATGGCTCTACCTGTGGCAGATGAC 1320
 Db 1272 TCTGGAGGAGTGGATCTGTGCGGACAAAGCAGATGGCTCTACCTGTGGCAGATGAC 1331
 QY 1321 AGAATGCTTTTGGGAGTGCATCAATGGAATCACATACACAGCAGATTTCTCAAGCAGGG 1380
 Db 1332 AGAATGCTTTTGGGAGTGCATCAATGGAATCACATACACAGCAGATTTCTCAAGCAGGG 1391
 QY 1381 CTTGTTTTGATACAGCTGTAATGTGCAATGGCCATGAACCTTAATGCCATTTCTTC 1440
 Db 1392 CTTGTTTTGATACAGCTGTAATGTGCAATGGCCATGAACCTTAATGCCATTTCTTC 1451
 QY 1441 AGAATTTCTGCACTCTCCTTTACTCTCAACAAAGTACTATCTCCCTTTAACTTA 1500
 Db 1452 AGAATTTCTGCACTCTCCTTTACTCTCAACAAAGTACTATCTCCCTTTAACTTA 1511
 QY 1501 TGCATTAATTTGGTAGCC 1519
 Db 1512 TGCATTAATTTGGTAGCC 1530

RESULT 5

BQ231098
 LOCUS AGENCOURT_7578138 NCI_CGAP_St1 Mus musculus cDNA clone
 DEFINITION IMAGE:6051580 5', mRNA sequence. EST 02-MAY-2002

BQ231098
 VERSION BQ231098.1 GI:20412498

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 892)

NIH-MGC http://mgi.nci.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabs-r@mail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov
 Plate: ILAM13305 row: k column: 05
 High quality sequence stop: 772.

FEATURES

source

1..892
 Location/Qualifiers
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="IMAGE:6051580"
 /clone_lib="NCI_CGAP_St1"
 /lab_host="DH10B (TI-resistant)"
 /note="Organ: stomach; Vector: pCMV-SPORT6; Site: 1: SalI;
 Site: 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
 Average insert size 1.77 kb. Library constructed by Life
 Technologies. Note: this is a NCI CGAP Library."

BASE COUNT 219 a 226 c 216 g 230 t 1 others

Query Match

Best Local Similarity 56.5%; Score 861.2; DB 14; Length 892;

Matches 869; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY	29	TGGCTCTTCTGCTGAATGCTCAGCTGGGGTCTGCCTACATCTGATATGCTATTTACCA 88
Db	1	TGGCTCTTCTGCTGAATGCTCAGCTGGGGTCTGCCTACATCTGATATGCTATTTACCA 60
QY	89	ACTGGGCCCCAGTATCGGCCAGGTCTGGGGAGCTTCAAGCCTGATGACATTAACCCCTGCC 148
Db	61	ACTGGGCCCCAGTATCGGCCAGGTCTGGGGAGCTTCAAGCCTGATGACATTAACCCCTGCC 120
QY	149	TGTGTACTACCTGATCTATGCTTGTCTGGGATGCAGAACATAGATCACCACCATAG 208
Db	121	TGTGTACTACCTGATCTATGCTTGTCTGGGATGCAGAACATAGATCACCACCATAG 180
QY	209	AATGGGAATGATGTACTCTCTATAAAGCTTTCATGACTTGAACACAGGAAACAGCAAC 268
Db	181	AATGGGAATGATGTACTCTCTATAAAGCTTTCATGACTTGAACACAGGAAACAGCAAC 240
QY	269	TGAAACCCCTCCCTGGCAATTTGGAGGCTTGAACCTTGTGAACCTGCTCTTTCACATGAG 328
Db	241	TGAAACCCCTCCCTGGCAATTTGGAGGCTTGAACCTTGTGAACCTGCTCTTTCACATGAG 300
QY	329	TTTCCACTTCTCAGAACCGCCAGACCTTCATTAACCTCAGTCATCAAAATTTCTGGGTCAGT 388
Db	301	TTTCCACTTCTCAGAACCGCCAGACCTTCATTAACCTCAGTCATCAAAATTTCTGGGTCAGT 360
QY	389	ATGGGTTTGTATGAGCTGGACCTGGACTGGGAATACCCAGCTCAGCTGGGAGCCCTCCCTC 448
Db	361	ATGGGTTTGTATGAGCTGGACCTGGACTGGGAATACCCAGCTCAGCTGGGAGCCCTCCCTC 420
QY	449	AGGACAAGCATCTTTCTACTGTCTTGGTGAAGGAAATGCGTGAAGCTTTTGTAGCAGGAGG 508
Db	421	AGGACAAGCATCTTTCTACTGTCTTGGTGAAGGAAATGCGTGAAGCTTTTGTAGCAGGAGG 480
QY	509	CTATTGAGAGCAACAGCCCGACACTGATGGTACTGCTGCTAGCTAGCTGGGAGATTTCCTA 568
Db	481	CTATTGAGAGCAACAGCCCGACACTGATGGTACTGCTGCTAGCTAGCTGGGAGATTTCCTA 540
QY	569	ACATCCAGCTGGCTATGAGATCCCTGAACCTTTCTAAGTACCTGGGATTTTCATCCATGCTCA 628
Db	541	ACATCCAGCTGGCTATGAGATCCCTGAACCTTTCTAAGTACCTGGGATTTTCATCCATGCTCA 600
QY	629	TGACATPATGACCTTCCATGGCTCCTGGGAGGCTTACACTGGGAGAAATAGTCTCTTTTACA 688
Db	601	TGACATPATGACCTTCCATGGCTCCTGGGAGGCTTACACTGGGAGAAATAGTCTCTTTTACA 660
QY	689	AATACCTTACTGAGACTGGTAGCAATGCTACCTCAATGATGGATTTATGATCAATGAATTT 748
Db	661	AATACCTTACTGAGACTGGTAGCAATGCTACCTCAATGATGGATTTATGATCAATGAATTT 720
QY	749	GGAAGAACAAATGGAGCCCGCAGCTGAGAACTCATTTGTTGGATTTCACAGAGTATGGACACA 808
Db	721	GGAAGAACAAATGGAGCCCGCAGCTGAGAACTCATTTGTTGGATTTCACAGAGTATGGACACA 780

found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

Plate: LLAM10816 row: 9 column: 09
High quality sequence stop: 824.

FEATURES

source

1. .838
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4912544"
/clone_lib="NCI_CGAP_SG2"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: salivary gland; Vector: pCMV-SPORT6; Site_1:
NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo
dT. Average insert size 1.3 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT 209 a 211 c 198 g 220 t

ORIGIN
Query Match 54.0%; Score 823.6; DB 12; Length 838;
Best Local Similarity 98.9%; Pred. No. 1.1e-242;
Matches 829; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

```
Oy 31 GCTCTTCGCTGAATGCTCAGCTGGGCTCGCTTACAAATCTCATATGCTATTTTCCAAAC 90
Db 1 GCTCTTCGCTGAATGCTCAGCTGGGCTCGCTTACAAATCTCATATGCTATTTTCCAAAC 60
Oy 91 TGGGCCAGTATCGCCAGGCTCTGGGAGCTTCAAGCTGTATGATTAACCCCTGCCGTG 150
Db 61 TGGGCCAGTATCGCCAGGCTCTGGGAGCTTCAAGCTGTATGATTAACCCCTGCCGTG 120
Oy 151 TGTACTACCTGATCTATGCTCTTGTGGGATGCAAGCAATGAGATCACCACCATAGAA 210
Db 121 TGTACTACCTGATCTATGCTCTTGTGGGATGCAAGCAATGAGATCACCACCATAGAA 180
Oy 211 TGGAAATGATGTTACTCTCTATAAAGCTTTCAATGACTTTGAAACAGGCAACCACTG 270
Db 181 TGGAAATGATGTTACTCTCTATAAAGCTTTCAATGACTTTGAAACAGGCAACCACTG 240
Oy 271 AAAACCCCTCGGCAATGAGGCTGGAACCTTGAAGCTGCTTTCACCTACCATGCTT 330
Db 241 AAAACCCCTCGGCAATGAGGCTGGAACCTTGAAGCTGCTTTCACCTACCATGCTT 300
Oy 331 TCCACTTCTCAGAACCGCAGACCTTCAATACCTCAGTCATCAAAATTTCTGGTCAGTAT 390
Db 301 TCCACTTCTCAGAACCGCAGACCTTCAATACCTCAGTCATCAAAATTTCTGGTCAGTAT 360
Oy 391 GGGTTTGTAGTACCTGGACCTGGATGGGAATACCCAGGCTCAGCTGGAGCCCTCCTCAG 450
Db 361 GGGTTTGTAGTACCTGGACCTGGATGGGAATACCCAGGCTCAGCTGGAGCCCTCCTCAG 420
Oy 451 GACAGCATCTTTCACCTGCTCTGTTGAAGGAATGCGTGAAGCTTTTGAGCAGAGGCT 510
Db 421 GACAGCATCTTTCACCTGCTCTGTTGAAGGAATGCGTGAAGCTTTTGAGCAGAGGCT 480
Oy 511 ATTGAGCAACAGGCCAGACCTGATGTTTACTGCTGTGCTGTAGCTGGGATTTTCCAAAC 570
Db 481 ATTGAGCAACAGGCCAGACCTGATGTTTACTGCTGTGCTGTAGCTGGGATTTTCCAAAC 540
Oy 571 ATCCAGGCTGGCTATGATATCCCTGAACTTTTCTAAGTACCTGGATTTTCCATCATGTCATG 630
Db 541 ATCCAGGCTGGCTATGATATCCCTGAACTTTTCTAAGTACCTGGATTTTCCATCATGTCATG 600
Oy 631 ACATATGACCTCCATGGCTCTGGGAGGCTACACTGGGAGATAGTCTCTTTTACAAA 690
Db 601 ACATATGACCTCCATGGCTCTGGGAGGCTACACTGGGAGATAGTCTCTTTTACAAA 660
Oy 691 TACCTTACTGAGACTGGTAGCAATGCCCTACCTCAATGTGGATTTATGTCATGAACTATTGG 750
Db 661 TACCTTACTGAGACTGGTAGCAATGCCCTACCTCAATGTGGATTTATGTCATGAACTATTGG 720
Oy 751 AAGAACAAATGAGGCCCGAGCTGAGAAAGCTCATTTGTTGGATTTCCAGATTTGGACACACC 810
Db 721 AAGAACAAATGAGGCCCGAGCTGAGAAAGCTCATTTGTTGGATTTCCAGATTTGGACACACC 780
```

Db 721 AAGAACAAATGAGGCCCGAGCTGAGAAAGCTCATTTGTTGGATTTCCAGATTTGGACACACC 780
Oy 811 TTCATCTCTGAGAAACCCCTCTGATATGAATGATGTCGCCCTTACCTCTGGTGATGCC 868
Db 781 TTCATCTCTGAGAAACCCCTCTGATATGAATGATGTCGCCCTTACCTCTGGTGATGCC 838

RESULT 9

LOCUS BG868741 837 bp mRNA linear EST 29-MAY-2001
DEFINITION 602787570F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4913800 5',
mRNA sequence.

ACCESSION BG868741
VERSION BG868741
KEYWORDS EST, GI:14219281
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 837)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.
CNA Library Preparation: Life Technologies, Inc.
CNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

Plate: LLAM10819 row: k column: 17
High quality sequence stop: 836.
Location/Qualifiers

FEATURES

source

1. .837
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4913800"
/clone_lib="NCI_CGAP_SG2"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: salivary gland; Vector: pCMV-SPORT6; Site_1:
NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo
dT. Average insert size 1.3 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT 212 a 211 c 197 g 217 t

ORIGIN
Query Match 52.6%; Score 801.4; DB 12; Length 837;
Best Local Similarity 97.5%; Pred. No. 7.8e-236;
Matches 814; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

```
Oy 1 ATGGCCAAAGCTACTTCTCGTCACAGGCTGGCTCTTCTGCTGAATGCTCAGCTGGGCT 60
Db 3 ATGGCCAAAGCTACTTCTCGTCACAGGCTGGCTCTTCTGCTGAATGCTCAGCTGGGCT 62
Oy 61 GCCTACATCTGATATGCTATTTTCCAACTGGGCCAGATATCGCCAGGCTCTGGGAGC 120
Db 63 GCCTACATCTGATATGCTATTTTCCAACTGGGCCAGATATCGCCAGGCTCTGGGAGC 122
Oy 121 TTCAAGCCTGATGACATTAACCCCTGCTGCTACTCAGCTGATCTATGCTCTTGTCTGG 180
Db 123 TTCAAGCCTGATGACATTAACCCCTGCTGCTACTCAGCTGATCTATGCTCTTGTCTGG 182
Oy 181 ATGCAGAAATAGATCACCACCATAGAAATGAATGATGTTTACTCTCTATTAAGCTTTC 240
Db 183 ATGCAGAAATAGATCACCACCATAGAAATGAATGATGTTTACTCTCTATTAAGCTTTC 242
Oy 241 AATGACTTGAATAACAGGACAGCAAACTGAAACCCCTCTGGCAATGGAGGCTGGAAC 300
Db 243 AATGACTTGAATAACAGGACAGCAAACTGAAACCCCTCTGGCAATGGAGGCTGGAAC 302
```

```
QY 301 TTTGGAACTGCTCTTTCACTACATGCTTTTCCACTTCTCAGAACGCCAGACCTTCATT 360
Db 303 TTTGGAACTGCTCTTTCACTACATGCTTTTCCACTTCTCAGAACGCCAGACCTTCATT 362
QY 361 ACCTCAGTCAATCAAAATTCCTGCTCAGTATGGGTTTTCATGACTGGACCTGGAGTGGAA 420
Db 363 ACCTCAGTCAATCAAAATTCCTGCTCAGTATGGGTTTTCATGACTGGACCTGGAGTGGAA 422
QY 421 TACCCAGGCTCAGCTGGAGCCCTCCTCAGGACAAGCATCTCTTCACTGTCTCTGGTGAAG 480
Db 423 TACCCAGGCTCAGCTGGAGCCCTCCTCAGGACAAGCATCTCTTCACTGTCTCTGGTGAAG 482
QY 481 GAAATGCTGAAGCTTTTGAGAGGAGGCTATTGAGAGCAACAGGCCAGACCTGATGGTT 540
Db 483 GAAATGCTGAAGCTTTTGAGAGGAGGCTATTGAGAGCAACAGGCCAGACCTGATGGTT 542
QY 541 ACTGCTGCTAGCTGCTGGGATTTCCAAACATCCAGCTGGCTATGAGATCCCTGAACTT 600
Db 543 ACTGCTGCTAGCTGCTGGGATTTCCAAACATCCAGCTGGCTATGAGATCCCTGAACTT 602
QY 601 TCTAAGTACCTGGATTTATCCATGTCATGACATATGACCTCCATGGCTCCTGGAGGGC 660
Db 603 TCTAAGTACCTGGATTTATCCATGTCATGACATATGACCTCCATGGCTCCTGGAGGGC 662
QY 661 TACACTGGGAGAAATAGTCTCTTTTACAAATACCTTACTGAGACTGGTAGCAATGCTAC 720
Db 663 TACACTGGGAGAAATAGTCTCTTTTACAAATACCTTACTGAGACTGGTAGCAATGCTAC 722
QY 721 CTCATGTGGATTTGTCATGAATATTGGAAGAACATGGAGCCCGCAGCTGAGAAGCTC 780
Db 723 CTCATGTGGATTTGTCATGAATATTGGAAGAACATGGAGCCCGCAGCTGAGAAGCTC 782
QY 781 ATTTGTTGGATTTCCAGAGTATGACACACCTTCCATCCTGAGAAACCCCTCTGATA 835
Db 783 ATTTGTTGGATTTCCAGAGTATGACACACCTTCCATCCTGAGAAACCCCTCTGATA 837
```

RESULT 10

```
BG872136 855 bp mRNA linear EST 29-MAY-2001
LOCUS 602792974F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4924003 5',
DEFINITION mRNA sequence.
```

```
ACCESSION BG872136
VERSION BG872136.1 GI:14222676
KEYWORDS EST.
```

```
SOURCE house mouse.
```

ORGANISM

```
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
```

```
1 (bases 1 to 855)
```

```
NIH-MGC http://mgi.nci.nih.gov/.
```

```
National Institutes of Health, Mammalian Gene Collection (MGC)
```

```
Unpublished (1999)
```

```
Contact: Robert Strausberg, Ph.D.
```

```
Email: cgabbs-r@mail.nih.gov
```

```
Tissue Procurement: Jeffrey E. Green, M.D.
```

```
cDNA Library Preparation: Life Technologies, Inc.
```

```
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
```

```
DNA Sequencing by: Incyte Genomics, Inc.
```

```
Clone Distribution: MGC clone distribution information can be
```

```
found through the I.M.A.G.E. Consortium/LLNL at:
```

```
http://image.llnl.gov
```

```
Plate: LLAM10846 row: d column: 20
```

```
High quality sequence stop: 836.
```

```
Location/Qualifiers
```

FEATURES

source

```
1..855
```

```
/organism="Mus musculus"
```

```
/strain="FVB/N"
```

```
/db_xref="taxon:10090"
```

```
/clone="IMAGE:4924003"
```

```
/clone_lib="NCI_CGAP_SG2"
```

```
/lab_host="PH10B (T1 phage-resistant)"
```

```
/note="Organ: salivary gland; Vector: pCMV-SPORT6; Site_1:"
```

RESULT 11

BG869655

LOCUS

DEFINITION

ACCESSION

```
602789716F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4921203 5',
mRNA sequence.
BG869655 836 bp mRNA linear EST 29-MAY-2001
BG869655
```

```
NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo
dt. Average insert size 1.3 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
```

```
BASE COUNT 216 a 218 c 201 g 220 t
```

ORIGIN

```
Query Match 52.5%; Score 801.2; DB 12; Length 855;
Best Local Similarity 99.4%; Pred. No. 9.2e-236;
Matches 825; Conservative 0; Mismatches 3; Indels 2; Gaps 2;
QY 1 ATGGCCCAAGCTACTTTCCTGCTCAGAGTCTGGCTTCTGCTGATGCTCAGCTGAGGCTCT 60
Db 1 ATGGCCCAAGCTACTTTCCTGCTCAGAGTCTGGCTTCTGCTGATGCTCAGCTGAGGCTCT 60
QY 61 GCCTACAATCTGATATGCTATTTCACCAACTGGGCCAGTATCGGCCAGGCTCTGGGAGC 120
Db 61 GCCTACAATCTGATATGCTATTTCACCAACTGGGCCAGTATCGGCCAGGCTCTGGGAGC 120
QY 121 TTCAAGCCTGATGACATTAACCCCTGCTGTACTACCTGATCTTACTCTCTATAAAGCTTTC 180
Db 121 TTCAAGCCTGATGACATTAACCCCTGCTGTACTACCTGATCTTACTCTCTATAAAGCTTTC 180
QY 181 ATCAGAACAAATGAGATCACCACCATAGATGAATGATGTTACTCTCTATAAAGCTTTC 240
Db 181 ATCAGAACAAATGAGATCACCACCATAGATGAATGATGTTACTCTCTATAAAGCTTTC 240
QY 241 AATGACTTTGAAAACAGGAGAACAGCAAACTGAAAACCCCTCTGGCAATTTGGAGGCTGGAAC 300
Db 241 AATGACTTTGAAAACAGGAGAACAGCAAACTGAAAACCCCTCTGGCAATTTGGAGGCTGGAAC 300
QY 301 TTTGGAACTGCTCTTTCACTACATGCTTTTCCACTTCTCAGAACGCCAGACCTTCATT 360
Db 301 TTTGGAACTGCTCTTTCACTACATGCTTTTCCACTTCTCAGAACGCCAGACCTTCATT 360
QY 361 ACCTCAGTCAATCAAAATTCCTGCTCAGTATGGGTTTTCATGACTGGACCTGGAGTGGAA 420
Db 361 ACCTCAGTCAATCAAAATTCCTGCTCAGTATGGGTTTTCATGACTGGACCTGGAGTGGAA 420
QY 421 TACCCAGGCTCAGCTGGAGCCCTCCTCAGGACAAGCATCTTTCACCTGTCTCTGGTGAAG 480
Db 421 TACCCAGGCTCAGCTGGAGCCCTCCTCAGGACAAGCATCTTTCACCTGTCTCTGGTGAAG 480
QY 481 GAAATGCTGAAGCTTTTGGAGAGGAGGCTATTGAGAGCAACAGGCCAGACCTGATGGTT 540
Db 481 GAAATGCTGAAGCTTTTGGAGAGGAGGCTATTGAGAGCAACAGGCCAGACCTGATGGTT 540
QY 541 ACTGCTGCTAGCTGCTGGGATTTCCAAACATCCAGCTGGCTATGAGATCCCTGAACTT 600
Db 541 ACTGCTGCTAGCTGCTGGGATTTCCAAACATCCAGCTGGCTATGAGATCCCTGAACTT 600
QY 601 TCTAAGTACCTGGATTTTCATCCATGTCATGACATATGACCTCCATGGCTCCTGGAGGGC 660
Db 601 TCTAAGTACCTGGATTTTCATCCATGTCATGACATATGACCTCCATGGCTCCTGGAGGGC 660
QY 661 TACACTGGGAGAAATAGTCTCTTTTACAAATACCTTACTGAGACTGGTAGCAATGCTTAC 720
Db 661 TACACTGGGAGAAATAGTCTCTTTTACAAATACCTTACTGAGACTGGTAGCAATGCTTAC 720
QY 721 CTCATGTGAG-ATTATGTCATGAATAT-TGGGAAGAACAAATGGAGCCCGCAGCTGAGAAGC 778
Db 721 CTCATGTGAGCATTTATGATGAATCTCTGGAAGAACAAATGGAGCCCGCAGCTGAGAAGC 780
QY 779 TCATTGTTGGATTTCCAGAGTATGGACACACCTTTCATCTCTGAGAAACCC 828
Db 781 TCATTGTTGGATTTCCAGAGTATGGACACACCTTTCATCTCTGAGAAACCC 830
```

RESULT 11

BG869655

LOCUS

DEFINITION

ACCESSION


```

VERSION  BG869655.1  GI:14220195
KEYWORDS  EST.
SOURCE    house mouse.
ORGANISM  Mus musculus
REFERENCE  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS   NIH-MGC http://mgi.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
          Email: cnapbs-r@mail.nih.gov
          cDNA Library Preparation: Life Technologies, Inc.
          DNA Sequencing by: Incyte Genomics, Inc.
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LLAM10838 row: p column: 04
          High quality sequence stop: 824.
          Location/Qualifiers
            1..836
              /organism="Mus musculus"
              /strain="FVB/N"
              /db_xref="taxon:10090"
              /clone="IMAGE:4921203"
              /clone_lib="NCI_CGAP_SG2"
              /lab_host="DH10B (TI phage-resistant)"
              /note="Organ: salivary gland; Vector: pCMV-SPORT6; Site_1:
              NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo
              dt. Average insert size 1.3 kb. Constructed by Life
              Technologies. Note: this is a NCI_CGAP Library."
          BASE COUNT  213 a 207 c 196 g 220 t
          ORIGIN
            Query Match      52.2%; Score 796.2; DB 12; Length 836;
            Best Local Similarity 98.2%; Pred. No. 3.2e-234;
            Matches 816; Conservative 0; Mismatches 13; Indels 2; Gaps 1;
          QY  1  ATGGCCAGCTACTCTCGTCACAGGCTGGCTCTTCTGCTGAATGCTCAGCTGGGGTCT 60
          DB  1  ATGGCCAGCTACTCTCGTCACAGGCTGGCTCTTCTGCTGAATGCTCAGCTGGGGTCT 60
          QY  61  GCCTACATCTGATATGCTATTTCACCAACTGGGCCAGTAGTCGCCAGGCTCGGGAGC 120
          DB  61  GCCTACATCTGATATGCTATTTCACCAACTGGGCCAGTAGTCGCCAGGCTCGGGAGC 120
          QY  121  TTCAAGCCTGATGACATTAAACCCCTGCTGTGTACTACCTGATGCTATGCTGCTGGG 180
          DB  121  TTCAAGCCTGATGACATTAAACCCCTGCTGTGTACTACCTGATGCTATGCTGCTGGG 180
          QY  181  ATGCAGACATGATGATCACCACCATAGATGGAATGATGTTACTCTCTATAAGCTTTC 240
          DB  181  ATGCAGACATGATGATCACCACCATAGATGGAATGATGTTACTCTCTATAAGCTTTC 240
          QY  241  AATGACTTGAACAGCAAGCAAACTGAACCCCTCGGCAATGGAGCTGGAGC 300
          DB  241  AATGACTTGAACAGCAAGCAAACTGAACCCCTCGGCAATGGAGCTGGAGC 300
          QY  301  TTTGGAAGCTGCTCTTTTCACTACCATGTTTCCACTTCTCAGAACCGGACACCTTCATT 360
          DB  301  TTTGGAAGCTGCTCTTTTCACTACCATGTTTCCACTTCTCAGAACCGGACACCTTCATT 360
          QY  361  ACCTCAGTCATCAAAATTTCTGGTCAGTATGGGTTTGATGGACTGGACCTGGAGCTGGAA 420
          DB  361  ACCTCAGTCATCAAAATTTCTGGTCAGTATGGGTTTGATGGACTGGACCTGGAGCTGGAA 420
          QY  421  TACCAGGCTCAGCTGGAGGCTCCTCAGGACAGCATCTCTTCACTGTCTCTGTGTGAAG 480
          DB  421  TACCAGGCTCAGCTGGAGGCTCCTCAGGACAGCATCTCTTCACTGTCTCTGTGTGAAG 480
          QY  481  GAAATGCGTGAAGCTTTTGGAGAGGAGGCTATTGAGAGCAACAGGCCCCAGACTGATGTT 540
          DB  481  GAAATGCGTGAAGCTTTTGGAGAGGAGGCTATTGAGAGCAACAGGCCCCAGACTGATGTT 540

          RESULT 12
          LOCUS  BG871468
          DEFINITION  602790678P1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4921890 5',
          mRNA sequence.
          ACCESSION  BG871468
          VERSION  BG871468.1  GI:14222008
          KEYWORDS  EST.
          SOURCE    house mouse.
          ORGANISM  Mus musculus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
          NIH-MGC http://mgi.nci.nih.gov/.
          National Institutes of Health, Mammalian Gene Collection (MGC)
          Unpublished (1999)
          Contact: Robert Strausberg, Ph.D.
          Email: cnapbs-r@mail.nih.gov
          Tissue Procurement: Jeffrey E. Green, M.D.
          cDNA Library Preparation: Life Technologies, Inc.
          DNA Sequencing by: Incyte Genomics, Inc.
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LLAM10840 row: 1 column: 19
          High quality sequence stop: 845.
          Location/Qualifiers
            1..880
              /organism="Mus musculus"
              /strain="FVB/N"
              /db_xref="taxon:10090"
              /clone="IMAGE:4921890"
              /clone_lib="NCI_CGAP_SG2"
              /lab_host="DH10B (TI phage-resistant)"
              /note="Organ: salivary gland; Vector: pCMV-SPORT6; Site_1:
              NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo
              dt. Average insert size 1.3 kb. Constructed by Life
              Technologies. Note: this is a NCI_CGAP Library."
          BASE COUNT  218 a 220 c 211 g 231 t
          ORIGIN
            Query Match      52.1%; Score 794.6; DB 12; Length 880;
            Best Local Similarity 97.3%; Pred. No. 1e-233;
            Matches 851; Conservative 0; Mismatches 19; Indels 5; Gaps 4;
          QY  4  GCCAAGCTACTTCTCGTCACAGGCTGGCTCTTCTGCTGAATGCTCAGCTGGGGTCTGCC 63
          DB  1  GCCAAGCTACTTCTCGTCACAGGCTGGCTCTTCTGCTGAATGCTCAGCTGGGGTCTGCC 60

```

Qy 64 TACAATCTGATATGCTATTTTCAACAACCTGGGCCAGTATCGGCCAGGTCTGGGAGCTTC 123
Db 61 TACAATCTGATATGCTATTTTCAACAACCTGGGCCAGTATCGGCCAGGTCTGGGAGCTTC 120
Qy 124 AAGCCTGATGACATTAACCCCTGGCTGTGTACTACCTGATCTATGCTTGTGGGATG 183
Db 121 AAGCCTGATGACATTAACCCCTGGCTGTGTACTACCTGATCTATGCTTGTGGGATG 180
Qy 184 CAGAACAATGAGATCACCACATAGAAATGAATGATTTACTCTCTATAAAGCTTCAAT 243
Db 181 CAGAACAATGAGATCACCACATAGAAATGAATGATTTACTCTCTATAAAGCTTCAAT 240
Qy 244 GACTTGAANAACAGGACAGCAAACTGAAAACCCCTCTGCAATTTGGAGGCTGGAATTT 303
Db 241 GACTTGAANAACAGGACAGCAAACTGAAAACCCCTCTGCAATTTGGAGGCTGGAATTT 300
Qy 304 GGAAGCTCTCTTCACTACCATGTTTCCACTTCTCAGAACCCGACCTCATTTACC 363
Db 301 GGAAGCTCTCTTCACTACCATGTTTCCACTTCTCAGAACCCGACCTCATTTACC 360
Qy 364 TCAGTATCAAAATTTCTGCTGATGATGGTTTGTGAGTGGACTGGACTGGGAATAC 423
Db 361 TCAGTATCAAAATTTCTGCTGATGATGGTTTGTGAGTGGACTGGACTGGGAATAC 420
Qy 424 CAGAGCTACGTGGGAGCCCTCTCAGGACAAGCATCTCTCACTGTCTTGGTGAAGGAA 483
Db 421 CAGAGCTACGTGGGAGCCCTCTCAGGACAAGCATCTCTCACTGTCTTGGTGAAGGAA 480
Qy 484 ATGCGTCAAGCTTTGAGCAGGAGGCTATTGAGAGCAACAGGCCAGACTGATGGTACT 543
Db 481 ATGCGTCAAGCTTTGAGCAGGAGGCTATTGAGAGCAACAGGCCAGACTGATGGTACT 540
Qy 544 GCTGCTAGCTGGTGGATTTCCAAATCAGCGTGGCTATGAGATCCCTGAACTTCT 603
Db 541 GCTGCTAGCTGGTGGATTTCCAAATCAGCGTGGCTATGAGATCCCTGAACTTCT 600
Qy 604 AAGTACCTGATTTATCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 663
Db 601 AAGTACCTGATTTATCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
Qy 664 ACTGGGAGAAATAGTCTCTTTTCAAAATACCTTACCTGAGACTGGTACGATGCTACCTC 723
Db 661 ACTGGGAGAAATAGTCTCTTTTCAAAATACCTTACCTGAGACTGGTACGATGCTACCTC 720
Qy 724 AATGTGATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 783
Db 721 AATGTGATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 779
Qy 784 GTTGGATTTCCAGAGTAT--GGACACCTTCACTGAGAACCCCTCTGATAATGGAA 841
Db 780 GCTGGATTTCAAGAGTATTTGGTGAACACCTTCACTGAGAA--CCCTCTGATCTGG-A 837
Qy 842 TTGGTGCCTTACCTCTGTTGATGGCCCTGCTGGC 876
Db 838 TTGGTGCCTTACCTCTGTTGATGGCCCTGCTGGC 872

RESULT 13
BG869958
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE

894 bp mRNA linear EST 29-MAY-2001
602789826F1 NCI_CGAP_SG2 Mus musculus cdna clone IMAGE:4920918 5',
mRNA sequence.
BG869958
BG869958.1 GI:14220498
EST.
house mouse.
Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 894)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL
COMMENT

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-f@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLNL0838 row: d column: 07
High quality sequence stop: 805.
Location/Qualifiers
1. .894
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4920918"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: salivary gland; Vector: pCMV-SPORT6; Site: 1;
NotI: Site 2; SalI: Cloned unidirectionally. Primer: Oligo
dT. Average insert size 1.3 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

FEATURES
source

BASE COUNT 221 a 225 c 214 g 234 t
ORIGIN
Query Match 51.8%; Score 789.8; DB 12; Length 894;
Best Local Similarity 96.5%; Pred. No. 3.2e-232;
Matches 850; Conservative 0; Mismatches 27; Indels 4; Gaps 4;
Qy 1 ATGCCCAAGCTACTTCTCGTCACAGGTCTGGCTTCTTCTGCTGAATGCTCAGCTGGGCTCT 60
Db 6 ATGCCCAAGCTACTTCTCGAAACAGCTCTGGCTTCTTCTGCTGAATGCTCAGCTGGGCTCT 65
Qy 61 GCCTACATCTGATGATGCTATTTCACCAACTGGGCCAGTATCGGCCAGGTCTGGGAGC 120
Db 66 GCCTACATCTGATGATGCTATTTCACCAACTGGGCCAGTATCGGCCAGGTCTGGGAGC 125
Qy 121 TTCAAGCCCTGATGACATTAACCCCTGCTGTGTACTCAGCTGATGCTGCTTGTCTGGG 180
Db 126 TTCAAGCCCTGATGACATTAACCCCTGCTGTGTACTCAGCTGATGCTGCTTGTCTGGG 185
Qy 181 ATGCAAGCAATGAGATCACCACATAGAAATGGAATGATGTTACTCTCTATAAAGCTTTC 240
Db 186 ATGCAAGCAATGAGATCACCACATAGAAATGGAATGATGTTACTCTCTATAAAGCTTTC 245
Qy 241 AATGACTTGAANAACAGGACAGCAAACTGAAAACCCCTCTGCAATTTGGAGGCTGGAAC 300
Db 246 AATGACTTGAANAACAGGACAGCAAACTGAAAACCCCTCTGCAATTTGGAGGCTGGAAC 305
Qy 301 TTGGAACTGCTCTTTTCACTACCATGGTTTCCACTTCTCAGAACCCGACAGCTTCAAT 360
Db 306 TTGGAACTGCTCTTTTCACTACCATGGTTTCCACTTCTCAGAACCCGACAGCTTCAAT 365
Qy 361 ACCTCAGTCATCAAAATTTCTGGTCTAGTATGGTTTGTGAGTGGACTGACCTGGAGAA 420
Db 366 ACCTCAGTCATCAAAATTTCTGGTCTAGTATGGTTTGTGAGTGGACTGACCTGGAGAA 425
Qy 421 TACCCAGGCTCAGTGGGAGCCCTCTCAGGACAAGCATCTCTTCACTCTCTGGTGAAG 480
Db 426 TACCCAGGCTCAGTGGGAGCCCTCTCAGGACAAGCATCTCTTCACTCTCTGGTGAAG 485
Qy 481 GAAATCGGTGAAGCTTTTTCAGCAGGAGGCTATTGAGCAACAAGGCCAGACTGATGGTT 540
Db 486 GAAATCGGTGAAGCTTTTTCAGCAGGAGGCTATTGAGCAACAAGGCCAGACTGATGGTT 545
Qy 541 ACTGCTGCTGATGCTGGGATTTCCAAATCCAGGCTGGCTATGAGATCCCTGAACTT 600
Db 546 ACTGCTGCTGATGCTGGGATTTCCAAATCCAGGCTGGCTATGAGATCCCTGAACTT 605
Qy 601 TCTAAGTACTCGATTTTCATCCATGATGATGATGATGATGATGATGATGATGATGATGAT 660

Db	606	TCTAAGTACCTGGATTTCATCATGTCATGACATATGACCTCCATGGCTCCTGGAGGC	665
Qy	661	TACACTGGG-GAGATAGTCCCTTTTACAAATACCTTAC-TGAGACTGGTAGCAATGGCT	718
Db	666	TACACTGGGAGATAGTCCCTTTTACAAATACCTTACCTTGGAGCTGGTAGCAATGGCT	725
Qy	719	ACCTCAATGTGGATTATGTCATGAAC-TATTGGAGAAACAATGGAGCCCCAGCTGAGAAG	777
Db	726	ACCTCAATGTGGATTATGTCATGAAC-TATTGGAGAAACAATGGAGCCCCAGCTTGAAG	785
Qy	778	CTCATGTGGATTCCAGAGATGACACACCTTCATCTCCTGAGAAACCCCTCTCATAT	837
Db	786	GTCAATGTGGATTCCAGAGATGACACACCTTCCTCTCTGAGAAA-CCCTCTCATAT	844
Qy	838	GGAATGTGGCCCTACCTCTGCTGATGGCCCTGCTGGGCG	878
Db	845	GGAATGTGGCCCTACCTCTGCTGATGGCCCTTTGGGCCC	885
RESULT 14			
LOCUS	BG866775	847 bp	mRNA linear EST 29-MAY-2001
DEFINITION	60278605F1 NCI_CGAP_SG2	Mus musculus	cdna clone IMAGE:4912608 5', mRNA sequence.
ACCESSION	BG866775		
VERSION	1		
KEYWORDS	EST.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
TITLE	NIH-MGC http://mgc.nci.nih.gov/		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999)		
Contact:	Robert Strausberg, Ph.D.		
Email:	cgapbs-remail.nih.gov		
Tissue Procurement:	Jeffrey E. Green, M.D.		
CDNA Library Preparation:	Life Technologies, Inc.		
CDNA Sequencing by:	Incyte Genomics, Inc.		
Clone distribution:	MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov		
Plate:	LLAM10816	row: j	column: 01
High quality sequence stop:	774.		
Location/Qualifiers			
1. .847			
/organism="Mus musculus"			
/strain="FVB/N"			
/db_xref="taxon:10090"			
/clone="IMAGE:4912608"			
/clone_lib="NCI_CGAP_SG2"			
/lab_host="DH10B (X1 phage-resistant)"			
/note="Organ: salivary gland; Vector: pCMV-SPORT6; Site: 1: NotI; Site: 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.3 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP Library."			
BASE COUNT	214 a	211 c	201 g 221 t
ORIGIN			
Query Match	51.7%;	Score 788.4;	DB 12; Length 847;
Best Local Similarity	96.8%;	Pred. No. 8.2e-232;	
Matches	815;	Conservative 0;	Mismatches 26; Indels 1; Gaps 1;
Qy	1	ATGGCCAGCTACTTCTCTGTCACAGTCTGGCTTCTCTGCTGAATGCTCAGCTGGGTCT	60
Db	6	ATGGCCAGCTACTTCTCTGTCACAGTCTGGCTTCTCTGCTGAATGCTCAGCTGGGTCT	65
Qy	61	GCTCAATCTGATGCTATTCACCACTGGGCCAGTATCGGCCAGCTGGGAGC	120
Db	66	GCTCAATCTGATGCTATTCACCACTGGGCCAGTATCGGCCAGCTGGGAGT	125
RESULT 15			
LOCUS	BG866417	809 bp	mRNA linear EST 29-MAY-2001
DEFINITION	602785352F1 NCI_CGAP_SG2	Mus musculus	cdna clone IMAGE:4911421 5', mRNA sequence.
ACCESSION	BG866417		
VERSION	1		
KEYWORDS	EST.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
TITLE	NIH-MGC http://mgc.nci.nih.gov/		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999)		
Contact:	Robert Strausberg, Ph.D.		
Email:	cgapbs-remail.nih.gov		
Tissue Procurement:	Jeffrey E. Green, M.D.		
CDNA Library Preparation:	Life Technologies, Inc.		
CDNA Sequencing by:	Incyte Genomics, Inc.		
Clone distribution:	MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov		
Plate:	LLAM10816	row: j	column: 01
High quality sequence stop:	774.		
Location/Qualifiers			
1. .847			
/organism="Mus musculus"			
/strain="FVB/N"			
/db_xref="taxon:10090"			
/clone="IMAGE:4912608"			
/clone_lib="NCI_CGAP_SG2"			
/lab_host="DH10B (X1 phage-resistant)"			
/note="Organ: salivary gland; Vector: pCMV-SPORT6; Site: 1: NotI; Site: 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.3 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP Library."			
BASE COUNT	214 a	211 c	201 g 221 t
ORIGIN			
Qy	121	TTCAAGCCTGATGACATTAAACCCCTGCTGTACTACCTGATCTATGCTCTTGTGGG	180
Db	126	TTCAAGCCTGATGACATTAAACCCCTGCTGTACTACCTGATCTATGCTCTTGTGGG	185
Qy	181	ATGCAGACAATGAGATCACCACCATAGATGAATGATGTTTACTCTCTATAAAGCTTTC	240
Db	186	ATGCAGACAATGAGATCACCACCATAGATGAATGATGTTTACTCTCTATAAAGCTTTC	245
Qy	241	AATGACTTTGAAAAACAGGAACAGCAAACTGAAACCCCTCTGCGCAATTGGAGCGTGAAC	300
Db	246	AATGACTTTGAAAAACAGGAACAGCAAACTGAAACCCCTCTGCGCAATTGGAGCGTGAAC	305
Qy	301	TTTGAACCTGCTCTTTTCTCACTACCATGTTTCCACTTCTCAGAACCGCCAGACCTTCATT	360
Db	306	TTTGAACCTGCTCTTTTCTCACTACCATGTTTCCACTTCTCAGAACCGCCAGACCTTCATT	365
Qy	361	ACCTCAGTCATCAAAATTTCTGCTCAGTATGGTTTGTAGTGGACTGGACCTGACACTGGGAA	420
Db	366	ACCTCAGTCATCAAAATTTCTGCTCAGTATGGTTTGTAGTGGACTGGACCTGACACTGGGAA	425
Qy	421	TACCCAGGCTCAGTGGGAGCCCTCTCTCAGGACAAGCATCTCTTCACTGCTCTGCTGGAAG	480
Db	426	TACCCAGGCTCAGTGGGAGCCCTCTCTCAGGACAAGCATCTCTTCACTGCTCTGCTGGAAG	485
Qy	481	GAATGCTGTAAGCTTTTGGAGCAGGAGGCTATTGAGAGCAACAGGCCCAAGACTGATGGTT	540
Db	486	GAATGCTGTAAGCTTTTGGAGCAGGAGGCTATTGAGAGCAACAGGCCCAAGACTGATGGTT	545
Qy	541	ACTGCTGCTGATGCTGGGATTTTCCAAACATCCAGCTGGCTGATGAGATCCCTGACTT	600
Db	546	ACTGCTGCTGATGCTGGGATTTTCCAAACATCCAGCTGGCTGATGAGATCCCTGACTT	605
Qy	601	TCTAAGTACCTGATTTTCACTCCATGACATATGACCTCCATGGCTCTCTGGGAGGCG	660
Db	606	TCTAAGTACCTGATTTTCACTCCATGACATATGACCTCCATGGCTCTCTGGGAGGCG	665
Qy	661	TACACTGGGAGATAGTCTCTTTTACAAATACCTTACGAGCTGGTATGAGCAATGCTTAC	720
Db	666	TACACTGGGAGATAGTCTCTTTTACAAATACCTTACGAGCTGGTATGAGCAATGCTTAC	725
Qy	721	CTCAATGTGGATTATGTCATGAATGATGAGCAATGAGAGCAATGAGAGCTGAGAGCTC	780
Db	726	CTCAATGTGGATTATGTCATGAATGATGAGCAATGAGAGCAATGAGAGCTGAGAGCTC	784
Qy	781	ATTGTTGGATTCCAGAGATGTCACACACCTTCATCTCAGAAACCCCTCTGATATGGA	840
Db	785	ATTGTTGGATTCCAGAGATGTCACACACCTTCATCTCAGAAACCCCTCTGATATGGA	844
Qy	841	AT 842	
Db	845	TT 846	

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: InCyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L14M10813 row: h column: 14
High quality sequence stop: 804.
Location/Qualifiers
1. 809
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4911421"
/clone_lib="NCI_CGAP_SG2"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: salivary gland; Vector: pCMV-SPORT6; Site: 1;
NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo
dT. Average insert size 1.3 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 193 a 200 c 208 g 207 t 1 others
ORIGIN

Query Match 51.7%; Score 788; DB 12; Length 809;
Best Local Similarity 99.1%; Pred. No. 1.1e-231;
Matches 802; Conservative 0; Mismatches 6; Indels 1; Gaps 1;
Qy 273 AACCTCTCGCAATGGAGCTGGAACTTTGGAACCTGCTCTTCACTACCATGGTTTC 332
Db 1 AACCTCTCGCAATGGAGCTGGAACTTTGGAACCTGCTCTTCACTACCATGGTTTC 60
Qy 333 CACTTCTCAGAACCGCCAGACCTTCATTACTCAGTCATCAAAATTTCTGGCTCAGTATGG 392
Db 61 CACTTCTCAGAACCGCCAGACCTTCATTACTCAGTCATCAAAATTTCTGGCTCAGTATGG 120
Qy 393 GTTTGATGGACTGGACCTGGGAAATCCAGGCTCAGTGGGAGCCCTCCTCAGGA 452
Db 121 GTTTGATGGACTGGACCTGGGAAATCCAGGCTCAGTGGGAGCCCTCCTCAGGA 180
Qy 453 CAAGCATCTCTTCACTGCTCTGTTGAAGAAATCGTGAAGCTTTTGAGCAGGAGGCTAT 512
Db 181 CAAGCATCTCTTCACTGCTCTGTTGAAGAAATCGTGAAGCTTTTGAGCAGGAGGCTAT 240
Qy 513 TGAGAGCAACAGGCCAGACTGATGGTTACTGCTGTGTAGCTGGTGGGATTTCCAAACAT 572
Db 241 TGAGAGCAACAGGCCAGACTGATGGTTACTGCTGTGTAGCTGGTGGGATTTCCAAACAT 300
Qy 573 CCAGGCTGGCTATGAGATCCTGAACCTTCTTAAGTACCTGATTCATCCATGTCATGAC 632
Db 301 CCAGGCTGGCTATGAGATCCTGAACCTTCTTAAGTACCTGATTCATCCATGTCATGAC 360
Qy 633 ATATGACCTCCATGGCTCTGGAGGGCTACACTGGGGAGAAATAGTCTCTTTACAATA 692
Db 361 ATATGACCTCCATGGCTCTGGAGGGCTACACTGGGGAGAAATAGTCTCTTTACAATA 420
Qy 693 CCCTACTGAGACTGGTAGCAATGCGTACCTCAATGTTGGATATGTCATGAACCTATTGGAA 752
Db 421 CCCTACTGAGACTGGTAGCAATGCGTACCTCAATGTTGGATATGTCATGAACCTATTGGAA 480
Qy 753 GAACAATGGAGCCCGCCAGCTCAGAGCTCATTGTTGGATTCCTCAGAGATGATGACACACCTT 812
Db 481 GAACAATGGAGCCCGCCAGCTCAGAGCTCATTGTTGGATTCCTCAGAGATGATGACACACCTT 540
Qy 813 CATCTCTGAGAAACCCCTCTGATATGGAATTTGGTGGCCCTACCTCTGGTGGTGGCCCTGC 872
Db 541 CATCTCTGAGAAACCCCTCTGATATGGAATTTGGTGGCCCTACCTCTGGTGGTGGCCCTGC 600
Qy 873 TGGGGCTATACAGACAGCTGGTTCTGGGCCCTACTATGAGATTTGACACCTTTCTGAG 932
Db 601 TGGGGCTATACAGACAGCTGGTTCTGGGCCCTACTATGAGATTTGACACCTTTCTGAG 660
Qy 933 AAGTGGAGCCACTGAGTCTT-GGGATGCCCTCCCAAGAGTGGCCCTATGCCCTATAAGGCCA 991
Db 661 AAGTGGAGCCACTGAGTCTTGGGGATGGCTTGGGGATGGCTTGGGGATGGCTTGGGGATGG 720

Qy 992 ACAGTGGCTTGGCTATGACAATATCAAGAGCTTCAGTGTAAAGCTCAGTGGCTTAAGC 1051
Db 721 ACAGTGGCTTGGCTATGACAATATCAAGAGCTTCAGTGTAAAGCTCAGTGGCTTAAGC 780
Qy 1052 AGAACAATTTTGGAGGTGCCATGATCTGG 1080
Db 781 AGAACAATTTTGGAGGTGCCATGATCTCGG 809

Search completed: July 3, 2003, 08:40:36
Job time : 2289.02 secs